Reactive Molecular Dynamics: From Small Molecules in Gas Phase to Enzymatic Reactions

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Sebastian Brickel

aus Deutschland

Basel, 2019

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von:

Prof. Dr. Markus Meuwly

Prof. Dr. Anatole von Lilienfeld

Basel, den 21. Mai 2019

Prof. Dr. Martin Spiess

Dekan

Preface

The most beautiful experience we can have is the mysterious.
It is the fundamental emotion which stands at the cradle of
true art and true science.

- *Albert Einstein* -

The mind is not a vessel to be filled but a fire to be kindled.

- *Plutarch* -

Acknowledgements

This work would not have been possible without the guidance and help of several individuals who have contributed to the success of this study.

First of all, I wish to thank my supervisor Prof. Dr. Markus Meuwly for guidance and help throughout this PhD. Furthermore I am grateful to Prof. Dr. Anatole von Lilienfeld for agreeing to co-referee this thesis and Prof. Dr. Stefan Willitsch for chairing the defence.

I would like to thank all of my colleagues, present and past, for their advice and the good times we shared. Special thanks go to Dr. Florent Hédin, Dr. Maksym Soloviov, Dr. Debasish Koner and Oliver Unke for scientific support. The group was like a family with an incredible team spirit.

I am especially thankful to Dr. Debasish Koner, Uxia Rivero Gonzalez, Ali Razzak and Felix Faber for revising my thesis.

Finally I want to thank my family for their support through all this time, and of course to my fiancée without whom I would be nothing.

Contents

Abstract

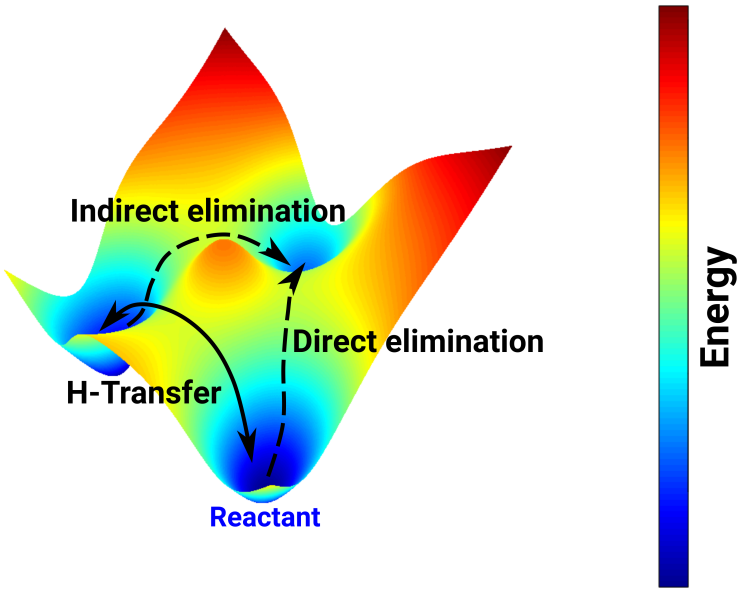
In this thesis, a comprehensive range of applications for the reactive molecular dynamics method Multi-Surface Adiabatic Reactive Molecular Dynamics (MS-ARMD) will be presented. My research investigates molecules calculated in gas phase as well as simulations in solution and in solvated enzymes. The motivation for each simulations will be highlighted, the parametrization of the energy functions explained and comparisons with other reactive Molecular Dynamics (MD) methods will be made.

Reactive molecular dynamics simulations allow the study of experimentally non amenable time- and length scales, providing essential insights into the mechanistic details of reactions. Knowledge drawn from such simulations allow i.a. the refinement of computational models and synthetic pathways, and finding new applications. The power of MS-ARMD lies in the calculation of converged reaction rates since thousands of individual trajectories can be run. This allows for the generation of statistically significant ensemble sizes for analysis such as quantitative characterization of final state distributions. This is usually not possible for conventional mixed quantum mechanics/classical mechanics (QM/MM) molecular dynamics or full *ab initio* molecular dynamics simulations due to the computational cost of the quantum calculations.

Simulations in condensed phase allow for direct comparison with experiments, which are often performed in (aqueous) solution or in the presence of a biologically relevant enzyme. Furthermore, those simulations provide additional insight into the mechanism of reactions with chemical and biochemical interest, beyond experimental findings. It was shown that MS-ARMD force fields (FF) parametrized in gas phase can be applied directly, without further modifications, to condensed phase simulations. This allows for quantitatively fitted FF to be used for simulations in solution and enzymes which is an additional advantage of MS-ARMD in comparison to other reactive molecular dynamics methods.

Part I.

Introduction



Begin - to begin is half the work, let half still remain; again begin this, and thou wilt have finished.

- *Marcus Aurelius* -

1. Motivation

1.1. Reactive Molecular Dynamics

Chemical sciences looks to understand the nature of chemical reactions and reactivity, including how fast a reaction occurs under given conditions. There are two major domains of chemical reaction dynamics, motion (or rearrangement) and time. Both domains can display large variations. For example, reactions in solution typically occur within seconds,[1–3] yet the fundamental processes (i.e. bond breaking and forming) occur on much shorter timescales (femto- to picoseconds).[4,5] This means that the majority of reaction time is spent translating the molecules and redistributing the energy among the vibrational modes until the system is energetically and coordinate wise primed for the actual reaction step.[5] Studying the redistribution of energy and bond formation is experimentally difficult due to the short-lived nature of these transitions.[5–8] Computational techniques coupled with chemical theory have become vital in understanding scientific phenomena of this kind.

Atomistic simulations need to have a meaningful description of the inter- and intramolecular interactions along the progression coordinate, i.e. from reactant(s) to the product(s).[5,8] In general, such simulations can be grouped in two classes: those who solve the electronic Schrödinger equation, called *ab initio*, and those who are based on empirical force fields for describing the potential energy surface (PES).[6,9] Moreover, there are hybrid methods which combine both, first principle methods and FFs.[5,10] The dimensionality of the PES depends on the number of atoms N. The PES is described by a $3N-6$ ($3N-5$ for linear molecules) dimensional hypersurface ($3N$ internal, 6 translational and rotational degrees of freedom).[11] As a consequence complexity of atomistic simulations drastically grows with the number of atoms in the molecule(s) involved in the studied reaction.[5]

Methods solving the electronic Schrödinger equation or approximations to it, often refereed to as

quantum mechanics (QM) methods, are versatile and powerful approaches, as the only requirement is a configuration of the molecule(s) involved. However, the lag of efficiency as well as aspects like convergence of the wave function, limit the usability of these approaches.[5] Contrary to that, empirical FFs provide fast and efficient simulations, but require extensive knowledge of the system studied for an accurate description of the energetics (i.e. for the parametrisation). Furthermore conventional FF have to be expanded, if one wants to describe bond breaking/ forming reactions. This is commonly achieved using an adiabatic picture for describing reactant and product states.[6] Furthermore the combination of both, quantum mechanics and molecular mechanics (MM), yields QM/MM methods. There the total system is partitioned into a "reactive" QM region and a surrounding MM region,[10] which acts as a modulator.[12] A still remaining question for QM/MM simulations is the size the QM region is required to have in order to obtain converged results.[5]

Reactive, empirical FF methods, such as EVB (Empirical Valence Bond),[13] MS-VALBOND (Multi-State VALence BOND)[14] and MS-ARMD (Multi-Surface Adiabatic Reactive Molecular Dynamics)[15] have been used for several different reaction dynamics in the past.[14,16–27] Aforesaid simulations are based on well defined energy functions, parameterized to a reference, such as QM calculations or experimental values. The most well known method, EVB, was originally developed to describe reactions in condensed phase and utilized valence bond theory. Valence bond theory, used to describe e.g. the bonding in H_2[28] and H_3,[29] was a nearly forsaken method, due to its alleged failures in comparison to the molecular orbital theory.[30] Valence bond theory was revived in order to include eviromental effects and advanced to EVB.[13] EVB uses predefined resonance structures (i.e. bonding arrangements[13]), which are described along the diagonal of a matrix used to define the reaction.[6] For an ionic bond cleavage reaction of type $X-Y \rightarrow X^- + Y^+$ this means that two resonance froms ($\psi_1 = X-Y$ and $\psi_2 = X^-Y^+$) are required to follow the reaction.[13] Therefore knowledge of the reaction system has to be obtained *a priori*. The coupling between these states is achieved by the off-diagonal elements, whose definition can vary.[6] In general a mathematical function defined along a reaction coordinate, like a Gaussian. The potential energy surface of the system is then described from mixing of the resonance states,[13] such that the lowest electronic eigenvalue is given by

$$\begin{vmatrix} H_{1,1} - E & H_{1,2} - ES_{1,2} \\ H_{2,1} - ES_{2,1} & H_{2,2} - E \end{vmatrix} = 0 \tag{1.1}$$

where $H_{i,j} = \epsilon_{i,j} = \langle\psi_i|H|\psi_j\rangle$ and $S_{i,j} = \langle\psi_i|\psi_j\rangle$. Within EVB $H_{i,j}$ are described by empirical functions that describe the resonance structures.[31] Similar to EVB, MS-VALBOND[14] describes the total energy of the system by a Hamiltonian, where the diagonal elements are the force fields for the connectivities. The off-diagonal elements are the coupling between those connectivities. Another approach that uses predefined states, i.e. knowledge of the system, for following reactions in time is MS-ARMD. This reactive molecular dynamics implementation was used for the study of adiabatic reactions related to, e.g. atmospheric chemistry,[15,26,27] organic synthesis,[24,32] and biochemistry.[6,33] MS-ARMD, which is based on mixing of different FFs in energy space, enables the simulation of a statistically significant number of trajectories which allows for extensive studies on a system. Additionally, it can be used to calculate converged reaction rates, for example in vibrationally excited reaction dynamics. MS-ARMD will be discussed in detail in section 2.3.

A different approach than the above mentioned empirical force field methods is based on the, nealry linear relation between bond order and bond length.[34] The ReaxFF method[35] describes the energy of a system as

$$E = E_{bond} + E_{over} + E_{under} + E_{val} + E_{pen} + E_{tors} + E_{conj} + E_{VdW} + E_{coul} \tag{1.2}$$

where E_{bond} is bond energy, calculated from the bond order (itself beeing calculated from the atomic distances) where the minimum 0 would corresponds to no bond and the maximum 3 to a triple bond. E_{over} and E_{under} are the over- and under-coordinatione energy, penalty terms for atoms where the bond order exceets or lacks the number of total bonds as descirbed in valence theory. E_{val} the valence angle terms and the corresponding penalty term E_{pen} (used for atoms with two double bonds). The conjugation enery E_{conj} (for harmonic systems), as well as the van der Waals (E_{VdW}) and Coulomb terms (E_{coul}). Initially developed to describe reactions of hydrocarbons,[35] ReaxFF has been extended to e.g. describe organocatalyzed reaction[36] and alkaly-metal carbon interactions (i.e. for energy applications).[37]

1.2. OH Overtone Stretch Dynamics

An interesting field of application for reactive molecular dynamics are photochemical reactions. In general two different photochemical reaction options are available. Both are of significance to the

chemistry in the tropo- and stratosphere in the Earth's atmosphere.[38] First the excitation of electronic states by ultraviolet (UV) light and second, the vibrational excitation of the ground state by less energetic photons.[39–41] The latter has been shown to be of importance, under atmospheric conditions, for the production of aerosols.[39,42,43] The inclusion of those vibrational overtone induced reaction pathways leads to improved atmospheric models.[40] Moreover, vibrational excitation is a promising method for developing mode and bond specific unimolecular reactions that differ from the thermal reaction pathways.[40,44–47] Those overtone induced reactions can be driven by excitations off different vibrational modes, mostly of X-H stretching overtones (X = C, O, S, N) due to their position in the spectra which lies close to the energy of photons in the atmosphere (in the energy range from 30-60 kcal/mol).[39] These X-H modes can be seen as weakly coupled local oscillators. They can be modeled anharmonically (i.e. through an anharmonic oscillators).[39,40] Examples of studied X-H stretch induced dissociation include CH_2OH,[48,49] (per)nitric acid,[50–57] pyruvic acid,[58–60] CF_3COOH[61] and sulfonic acid.[51,62]

For sulfonic acid the excitation of an OH-stretching vibration leads to photofragmentation into water and sulfur-trioxide.[25,62] This can, suggestively, also occure under atmospheric conditions.[42,63,64] In the atmosphere the reaction is attributed to the so called "Junge layer", at around 15 to 35 km of altitude, where the cycling between SO_2 and H_2SO4 participates to stratospheric ozone depletion.[65–68] After excitation of the local OH-stretch normal mode H_2SO_4 can follow two different reaction pathways: intramolecular H-transfer or water elimination. These reactions have been investigated using QM/MD simulations e.g. at the PM3 level of theory[55] and other computational methods[69–72,72,73,73–75] including MS-ARMD.[25]

For the study of H_2SO_4 with MS-ARMD a global PES for describing the reaction pathways was parameterized to several thousand MP2/6-311G++(2d,2p) reference electronic structures. The PES, describing H_2SO_4, H_2O, SO_3, and the van der Waals complex of water and sulphur-trioxide contained a total of 16 states. These states are describing the permutatively invariant states of the reactant and the product. The excitation of the OH-stretching vibration was achieved by scaling the instantaneous velocities along the normal mode. A Morse potential was used for describing the excitation energies of the OH-stretch. Several thousand independent trajectories with a simulation time of 1 ns resulted in 58%, 77%, and 80% of water elimination for the excitation with $\nu_{OH} = 4$, 5, and 6, respectively. The

final state analysis, predicts the expected findings of an experimental investigation. It was found that if H-transfer precedes the water elimination (called indirect) the excess energy distributes preferably into vibrational and rotational degrees of freedom. In contrast, direct water elimination (without prior H-transfer) leads to redistribution into translational energy. Such results, obtained from statistically significant sampling sizes, simulated for long reaction time highlight the advantages of FF based reactive MD over *ab initio* MD. Another advantage of methods like MS-ARMD is there transferability, for example from gas to condensed phase.

1.3. Condensed Phase Reactions

Since most chemical reactions, especially biochemical reactions, take place in solution, the effect of solvation is important. Solvation leads to a change in the shape of the PES and has therefore an influence on many processes during the dynamics.[76,77] For instance, the S_N2 reaction is known to show an anti-catalytic effect, i.e. an increase in its reaction barrier, when going from gas phase to aqueous solution (or any polar solvent).[78–84] The opposite is true for the Claisen rearrangement which undergoes a reduction in the reaction barrier in polar solvent compared to gas phase. This is due to transition state (TS) stabilization. In the case of the Claisen rearrangement the solvent effect is in part based on dipole-dipole interactions.[85–87] Additionally the solvent affects branching ratios[88] and energy flow into products due to frequent collisions of the solute with the solvent.[77]

Another large field of research in condensed phase covers reactions with biochemical background, like enzymatic reactions.[8,89,90] Insight into the mechanistic details at the atomic level can provide broad synthetical[91] and technological applications[92,93] due to the intense catalytic power of enzymes.[90,94,95] Computer simulations in this field are capable of determining the structure and nature of the TS, something that is experimentally difficult to accurately obtain.[12,90,94] Additionally, calculations can identify functional groups and interactions essential for the catalysis.

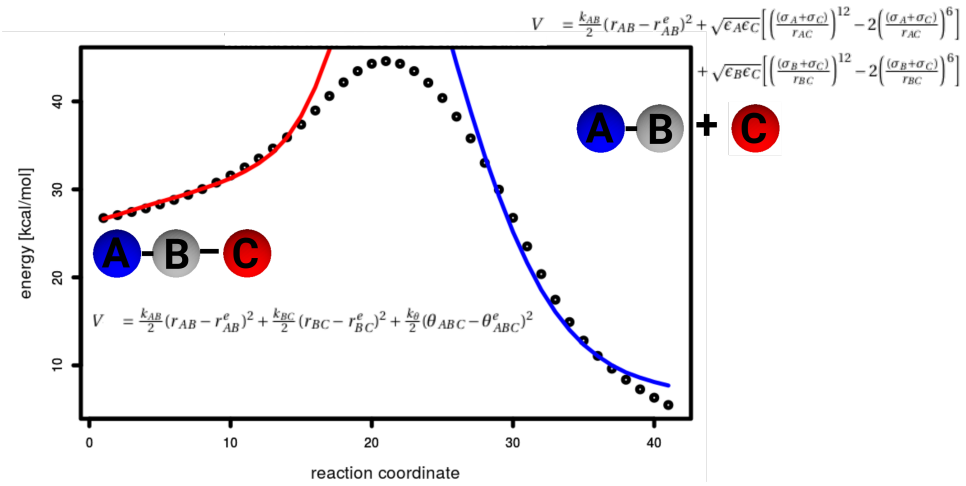
Simulations with biochemical structures usually start from X-ray data,[90] which may be solvated. In many cases the solvent does not actively participate in the reaction and can hence be excluded. In some cases truncated enzymes are studied. This can be done when there are metallic centers involved,

for example. Examples of such systems are Heme proteins,[10] like hemoglobin,[5,8,10,96] where the Fe center can bind several different small ligands like oxygen or nitrogen-oxide.[5,8,10,96] An important metric in computational simulations of enzymes is the free energy of activation, which can be compared to the barriers determined for the rate limiting step by experimental techniques.[12] For example, it was possible to show that the dicarboxylation reaction of hemoglobin is a fast, time limiting[97] and in fact the rate limiting step by comparing experiment and theory.[5,6]

In many biochemical applications,[90] but also generally for systems where the free energy surface (FES) is characterized by large barriers, extensive sampling of the conformational space is required.[90,98] A multitude of different methods have been developed to overcome those high barriers and explore the FES, such as thermodynamic integration,[99] Hamiltonian replica exchange MD[100] and umbrella sampling.[101]

Part II.

Methods



I begin to speak only when I'm certain what I'll say isn't better left unsaid.

- *Cato the Younger* -

2. Computational Methods

This chapter will summarize the most important methods utilized in the reactive molecular dynamics studies presented in this thesis, as well as other related methods.

2.1. Ab initio Methods

The fundamental postulate of quantum mechanics is the Schrödinger equation, which yields the wave function Ψ that completely describes every system. For many systems that are interesting for chemistry it is sufficient to solve the time-independent Schrödinger equation:[102]

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian operator and E the energy eigenvalue. Doing this without additional information taken from experiments leads to *ab initio* methods. The alternatives, i.e. including experimental data, are called semi-empirical methods. The coupling among nuclei and electronic motion in the Schrödinger equation is usually uncared for, which is called the Born-Oppenheimer approximation.[103] This approximation can be made because the electrons are much lighter than the nuclei, therefore the electrons move faster than the nuclei and the nuclei can be approximated as being stationary. This leads further to the neglection of the kinetic energy of the nuclei, as well as the assumption that the repulsion among the nuclei is constant.[104] This reduces the computational cost to (mostly) solving only the electronic Schrödinger equation.[105]

2.1.1. Hartree-Fock Theory

In order to solve many-electron systems additional approximations to the ones mentioned above are needed. The most pivotal method is the Hartree-Fock (HF) theory, where the many-electron problem is reduced to the interaction of one-electron with the averaged field of the remaining electrons.[104] Each electron is represented by an orbital and the overall wave function is described by a Slater determinant. HF is subject to the variational principle, which states that, within the limitation of using a single Slater determinant as the wave function, the HF orbitals yield the lowest energy[105] for equation 2.2.[104]

$$E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle = \sum_a \langle a|h|a\rangle + \frac{1}{2}\sum_{ab}\langle ab||ab\rangle = \sum_a \langle a|h|a\rangle + \frac{1}{2}\sum_{ab}[ab|ab] - [ab|ba] \tag{2.2}$$

where $\hat{H}$ is the electronic Hamiltonian, $\sum_a \langle a|h|a\rangle$ the 1-electron term and $\frac{1}{2}\sum_{ab}\langle ab||ab\rangle$ the 2-electron term. By using an iterative procedure a given orbital can be determined after calculating all remaining orbitals, which leads to self-consistent field (SCF) orbitals.[105]

2.1.2. Single-reference Perturbation Theory

Improvements over the simple HF method, which are generally grouped together as post HF methods, are e.g. perturbation theory (PT) methods. In PT the Hamiltonian operator is split into the zeroth-order or reference part ($\hat{H}_0$) and the perturbation ($\hat{H}'$). The later is assumed to be small when compared to $\hat{H}_0$.[104,105] The total Hamiltonian is written as

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}' \tag{2.3}$$

where λ is a parameter determining the strength of the perturbation. For $\lambda = 0$ one obtains the unperturbed Schrödinger equation. When the value of λ is increased to infinity in the perturbed Schrödinger equation ($\hat{H}\Psi = W\Psi$) it can be rewritten as a Taylor expansion, for both W and Ψ.[104,105]

$$W = \lambda^0 W_0 + \lambda^1 W_1 + \lambda^2 W_2 + \dots$$
$$\Psi = \lambda^0 \Psi_0 + \lambda^1 \Psi_1 + \lambda^2 \Psi_2 + \dots \tag{2.4}$$

PT converges slowly, compared to coupled cluster methods (another post HF method, most comonly extended to Single and Doubles(CCSD) or furthermore indcluding perturbative Triples (CCSD(T))), towards the full Configuration Interaction (CI) limit.[106] Where the full CI limit meaning the numerical exact solution to the electronic Schrödinger equation.[104]

Møller Plesset Perturbation Theory

Since its development in the 1930s the perturbation theory of Møller and Plesset[107] has become a popular PT method.[108] The zeroth-order part is described by the HF density, from which the perturbation term is constructed. Møller Plesset PT states that solving a one-body problem can be approximated by perturbatively solving the equation

$$F_\mu \psi = \kappa \psi \tag{2.5}$$

where F_μ is the one-electron operator, ψ are the eigenfunctions and κ the corresponding eigenvalues.[107] In their work the authors further state that, up to the second order, the HF density is accurate. In practical terms, as for all electron correlation methods, the integral transformation of two-electron integrals from atomic orbitals to molecular orbital basis is required.[108] For second-order Møller Plesset Perturbation Theory (MP2), among electron correlation methods the most efficient one, the transformation from two occupied (occ.;a, b) to two virtual (virt.;i, j) indices is needed. The MP2 energy for a closed shell system is then given as,[108]

$$E_{\mathrm{MP2}} = \sum_{\text{occ.}} \sum_{\text{virt.}} t^{\text{occ.}}_{\text{virt.}} (ai|bj) \tag{2.6}$$

where the integral $(ai|bj)$ is symmetrical and the amplitude $t^{\text{occ.}}_{\text{virt.}}$ is given by

$$t_{\text{virt.}}^{\text{occ.}} = \frac{2(ai|bj) - (bi|aj)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} \tag{2.7}$$

where ϵ gives the orbital energies[108]

MP2 scales as N^5 with respect to the number of orbitals N.[106] Many algorithms have been developed to improve the efficiency of MP2.[108]

2.1.3. Basis Set

Basis sets are used to describe molecular orbitals as a set of known, easy to calculate functions. Generally there are two types of basis functions used in computational chemistry: Slater Type Orbitals (STO) and Gaussian Type Orbitals (GTO). The two most commonly used families of basis sets are Pople[109] and Dunning[110] basis sets.[105] If the number of functions is chosen to fit exactly the number of electrons of the neutral atoms a, so called, minimum basis set is obtained. X-tuple (where X is D (double), T (triple), Q (quadruple)) the amount of valence basis sets leads to an improvement of the description and the class of VXZ (X = D, T, Q; in the nomenclature of Dunning) basis sets. Those can be further improved by including diffuse and polarization functions (p). Additional correlation consistency, accounting for the correlation energy of the valence electrons is included in correlation-consistent (prefix "cc") basis sets. Furthermore the inclusion of diffuse functions yields augmented basis sets (aug-cc-pVXT).

2.1.4. QM/MM

A popular application of *ab initio* quantum-mechanical methods in reactive molecular dynamics simulations is QM/MM. This is the combination of QM and molecular-mechanical (see Section 2.2) methods. QM/MM is applied broadly in, but not limited to, biomolecular studies,[10] such as enzymatic catalysis, where the mechanistic insight on the atomistic level provides a deeper understanding of the underlying chemistry.[90] The basic idea behind those hybrid methods is the partitioning of the system of interest into two parts, the inner, quantum mechanically described subsystem and the outer, molecular-mechanically described subsystem (see Figure 2.1). The relationship between the two

subsystems is described by the boundary region.[10] In the study of enzymatic catalysis the partitioning is usually between the (non-covalently bound) ligand and the way larger enzyme.[90] Many different methods for coupling the two subsystems, i.e. the description of the boundary region, exist. The most straightforward approach consists of calculating the interactions between the two subsystems using MM. Other, more advanced methods include polarization of the QM or the QM and MM subsystem, charge transfer[90] or link atoms.[10]

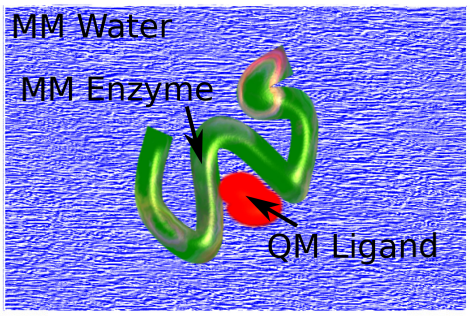


Figure 2.1.: Schematic representation of the QM/MM partition in an enzymatic study, solvated in water.

In terms of cost and efficiencies semi-empirical density functional tight binding (SCC-DFTB)[111,112] is among the best methods of choice for the QM part.[90] DFTB is based on the second-order expansion of the Kohn-Sham[113] total energy with respect to the fluctuation of the density with respect to a given reference. It self-consistently minimizes the total energy, taking into account the relaxation of the Mulliken population initial charge density.[111,112] The incorporation of DFTB into CHARMM[111] enables QM/MM calculations with ligands larger than a dozen heavy atoms.[96]

2.2. Force Fields

Among *ab initio* methods, which are only applicable to systems with a limited number of heavy atoms,[96] molecular mechanics methods are widely used to calculate the electronics yielding a PES. By applying Newton's laws and "ball and spring" models (where atoms are described as balls and bonds

between them as springs) these procedures are applicable to very large systems, such as enzymes. The general approximation applied in MM is based on the observation that structurally similar groups, i.e. functional groups, behave similarly in different molecules, which leads to so called atom types.[105] The potential energy functions applied in MM are called force fields.[10] The potential energy $U(\vec{R})$ ($\vec{R}$ being the Cartesian coordinates) is described as

$$U(\vec{R}) = \sum U_{\text{bonded}} + \sum U_{\text{nonbonded}} \tag{2.8}$$

where U_{bonded} consist of bond stretching, angle bending, and torsion terms and $U_{\text{nonbonded}}$ is generally split into van der Waals (vdW) and Coulomb interactions,[105] compare Figure 2.2.

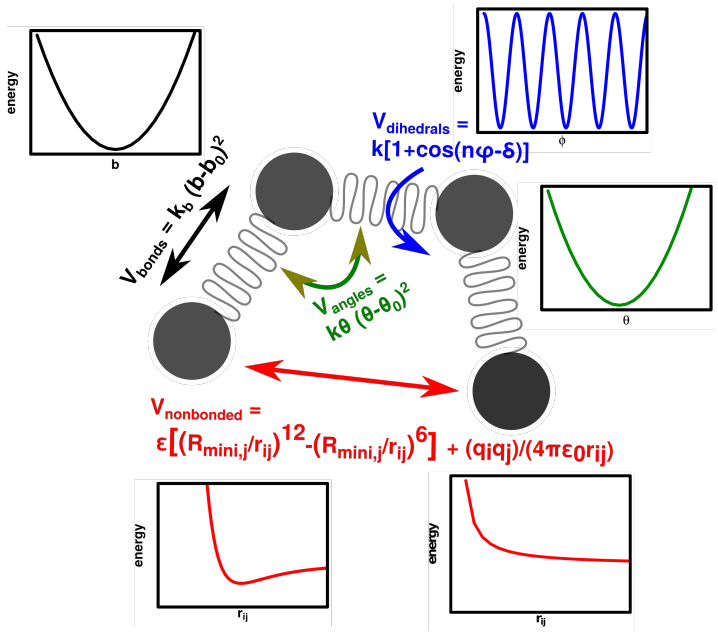


Figure 2.2.: Schematic representation of a four atomic molecule in ball and spring representation, including (harmonic) bond stretching, (harmonic) angle bending, dihedral/torsion, and non-bonded (Lennard-Jones (LJ) and Coulomb) interaction terms, as well as the corresponding energy function plots.

Commonly applied FF are, e.g. OPLS-AA,[114] Amber[115] and CHARMM[7] (see below). In general MM methods can not describe bond breaking or formation,[105] one possible approach to overcome this shortcoming is described below (see section 2.3). Furthermore, quantum effects such as tunneling are

not included im MM methods[116] but are assumed to be negligible in applications like biochemical systems.[9]

CHARMM Force Field

The CHARMM (Chemistry at HARvard Molecular Mechanics)[9] FF is a widely used, all atom force field with an emphasis on biomolecular systems, like proteins, peptides, and lipids. The form of the potential $U(\vec{R})$ is given by the bonded terms described by harmonic bonds (β) and angles (θ). Furthermore sinusoidal dihedrals (φ), improper dihedrals (ω) and Urey Bradley 1-3 terms (S). The nonbonded terms are given by electrostatic Coulomb (conventionally with fixed point charges) and van der Waals interactions (6-12 Lennard Jones potential).[9]

$$\begin{aligned} U(\vec{R}) = &\sum_{\text{bonds}} K_b(\beta-\beta_0)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \sum_{\text{Urey-Bradley}} K_{UB}(S-S_0)^2 \\ &+ \sum_{\text{dihedrals}} K_\varphi\big(1+cos(n\varphi-\delta)\big) + \sum_{\text{impropers}} K_\omega(\omega-\omega_0)^2 \\ &+ \sum_{\text{non-bonded pairs}} \Big[\epsilon_{ij}^{min}\Big[\big(\frac{R_{ij,min}}{r_{ij}}\big)^{12} - \big(\frac{R_{ij,min}}{r_{ij}}\big)^{6}\Big] + \frac{q_i q_j}{4\pi\epsilon_0\epsilon r_{ij}}\Big] \end{aligned} \tag{2.9}$$

The respective force constant is given by K, subscript 0 corresponds to the equilibrium value, n is the multiplicity of the dihedral and δ is the phase shift of the dihedral. This simple form of the potential gives a good compromise between accuracy and speed, as required for large biochemical systems.[9] For the recent CGenFF (CHARMM Generalized FF[117]) the focus of the parametrisation was put towards QM reference, in order to ease the parametrisation of biochemical ligands. The reference method of choice is MP2 with a Pople style basis set. The goal was to provide sufficiently accurate chemical building blocks for a broad range of chemical applications.[9] The continuous development of the CHARMM FF nowadays also include, among others, models for describing induced polarization[9] and fluctuating point charges.[118] Online tools, such as the CHARMM-GUI website,[119] have been developed to facilitate the usage of the CHARMM FF, especially the generation of new ligand force fields (through their ligand reader).[120]

2.3. Multi-Surface Adiabatic Reactive Molecular Dynamics

MS-ARMD[15] is a broadly applicable, energy conserving surface crossing algorithm, which in contrast to other reactive force field methods, is based on the potential energy of the studied system. MS-ARMD uses parameterized FF for the reactant and (multiple) product state(s), called surfaces, and, so called, GAPOs (GAussian × POlynomials) to describe the surface crossing region, i.e. the area around the transition state (TS). Since conventional force field terms are too rigid for bonds to elongate sufficiently long and for atoms to approach adequately close MS-ARMD allows to replace the harmonic bonds and 6-12 Lennard Jones potentials by Morse potentials

$$V_{\text{Morse}}(r) = D_e(1 - exp(-\alpha(r - r_e)^2)) \tag{2.10}$$

(with r the distance between the atoms, r_e the equilibrium bond distance, D_e the well depth, and α the width of the potential) and Generalized Lennard Jones potentials,

$$V_{\text{GLJ}}(r; n, m, \epsilon, r_{\text{min}}) = \frac{(n\epsilon)}{(m-n)}\left[\left(\frac{r_{\text{min}}}{r}\right)^m - \frac{m}{n}\left(\frac{r_{\text{min}}}{r}\right)^n\right] \tag{2.11}$$

respectively. Where the the well depth of the Generalized Lennard Jones potential is described by ϵ and the separation at the energy minimum by r_{min}. These two exchanges of the potential are necessary for correctly describing the potential energy along the reaction path and therefore required for those atoms involved in the bond braking and bond forming (see Figure 2.3).

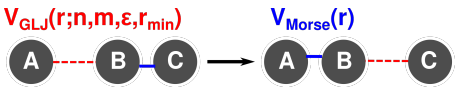


Figure 2.3.: Sketch of the application of Morse potentials and generalized Lennard Jones potentials for describing the bond breaking and bond forming interactions in MS-ARMD.

In order to combine the individual surfaces which describe the reactant and product state(s), energy dependent weights, mathematically described as coordinate dependent ($\vec{x}$: Cartesian coordinates), $w_i(\vec{x})$ for each surface $V_i(\vec{x})$ are assigned. The particular weights are calculated by renormalizing the

raw weights $w_{i,0}(\vec{x})$ by the following equation:

$$w_i(\vec{x}) = \frac{w_{i,0}(\vec{x})}{\sum_{j=1}^{n} w_{j,0}(\vec{x})} \tag{2.12}$$

The raw weights are evaluated by using the potential energy difference between the i-th surface and the minimum energy surface $V_{\min}(\vec{x})$ with an exponential decay parameter ΔV.

$$w_{i,0}(\vec{x}) = \exp\big(-\frac{V_i(\vec{x}) - V_{\min}(\vec{x})}{\Delta V}\big) \tag{2.13}$$

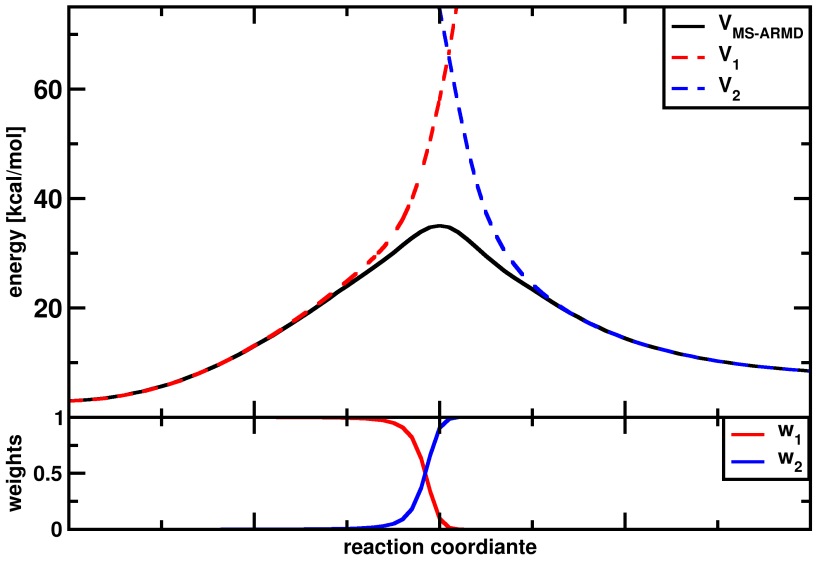


Figure 2.4.: Top panel) Example of MS-ARMD global, reactive surface construction based on two surfaces (V_1, red dashed; V_2, blue dashed). The multi-dimensional PES ($V_{\text{MS-ARMD}}$, black solid) matches the lowest-energy surface, except for the crossing region, where the weights (bottom panel, w_1, w_2) switch smoothly between the two surfaces. Reaction coordinate in arbitrary units, weights are unit less.

The resulting global surface is then given by

$$V_{\text{MS-ARMD}}(\vec{x}) = \sum_{i=1}^{n} w_i(\vec{x}) V_i(\vec{x}) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} [w_i(\vec{x}) + w_j(\vec{x})] \sum_{k=1}^{n_{ij}} \Delta V^{ij}_{\text{GAPO},k}(\vec{x}) \tag{2.14}$$

where

$$\Delta V^{ij}_{\text{GAPO},k}(\vec{x}) = exp\Big(- \frac{(\Delta V_{ij}(\vec{x}) - V^0_{ij,k})^2}{2\sigma^2_{ij,k}} \Big) \times \sum_{l=0}^{m_{ij,k}} a_{ij,kl} (\Delta V_{ij}(\vec{x}) - V^0_{ij,k})^l \tag{2.15}$$

gives the energy contribution of the GAPOs for the energy difference $\Delta V_{ij}(\vec{x}) = V_j(\vec{x}) - V_i(\vec{x})$. The center and standard deviation of the Gaussian function is given by $V^0_{ij,k}$ and $\sigma_{ij,k}$, respectively, while the polynomial of order $m_{ij,k}$ has the coefficient $a_{ij,kl}$. An example of a MS-ARMD global reactive surface is given in Figure 2.4.

Fitting Algorithms for Force Field Parametrisation

The parametrisation strategy for MS-ARMD applied throughout this work consists of an iterative procedure of generating ensembles of structures from a FF, evaluating the QM reference energies (e.g. from Gaussian09[121]) for those structures and fitting the FF parameters to reproduce the QM reference energies (within chemical accuracy of $\approx$ 1 kcal/mol). As written above (see subsection 2.2), CGenFF already heavily depends on MP2 reference calculations and hence the discrepancy between the CGenFF and QM reference methods is small when MP2 is chosen as reference. This small discrepency allows the usage of a nonlinear downhill Simplex algorithm.[122] For the parametrisation of a specific FF for each molecular reaction studied, CGenFF provides a good initial guess. Downhill simplex, a robust algorithm which falls into the class of gradient decent algorithms, uses simplices, polyhedrons containing $N + 1$ points in N dimensional space. In each iteration step the method will replace the vertices of the polyhedron with the highest functional value (differences to reference) by a new one using, so called reflection (see Figure 2.5). Reflection means that the aforementioned vertices are reflected in the opposite direction. This step is followed by the expansion (see Figure 2.5), where the change in the vertices is continued in the same direction as in the reflection step, if the new simplex gave a better value than the original one. If this is not the case the method performs contraction (see Figure 2.5).

There the newly 'reflected' point is exchanged with the worst one and a new simplex is constructed. Since no derivatives are needed in this procedure, no gradient based termination criterion is used. Instead the algorithm will terminate if no further improvement can be found (within some predefined tolerance value). Since the construction of the initial simplex depends on the initial parameters and their quality, a good guess to the minimum has to be provided.

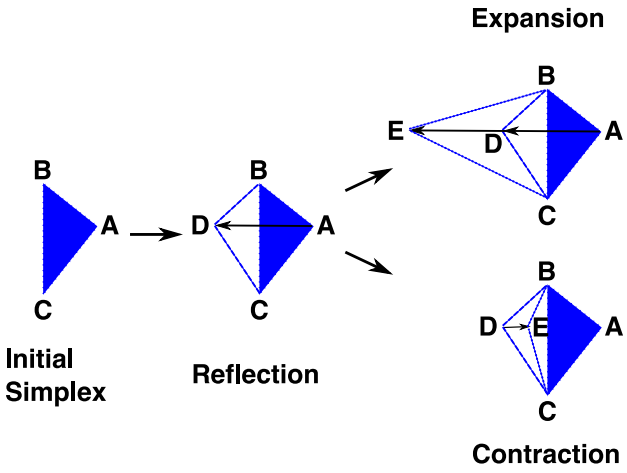


Figure 2.5.: Types of simplex moves. From left to right: initial simplex; reflection step, with the reflected point D; top) expansion with new point E; bottom) contraction with contracted point E.

For the FFs of the individual molecules this procedure leads to convergence of the FF with respect to the QM reference methods within a few iterations (usually three to four). For the parametrisation of the GAPOs in MS-ARMD the fitting strategy has to be changed since the GAPOs are highly multidimensional, abstract objects in energy difference space, for which a good initial guess is not available. Therefore a Genetic Algorithm (GA) was applied to reproduce the potential energy along an Intrinsic Reaction Coordinate (IRC) or Minimum Energy Path (MEP), i.e. the connection between reactant and product following the lowest potential energy path, based on the parameterized FFs involved. Genetic algorithms are based on the evolution of the fittest and deal with populations of possibilities whose fitness are evaluated, crossed with others (according to their fitness) and mutated. Over several iterations the algorithm converges to a minimum.[123,124] The applied algorithm was based on differential evolution,[125] a method which explores low value regions of a given function. The algorithm was adopted from a previous work[126] using a given, large number of generations and

several dozen points in the population (dependent on the number of parameters). A flowchart for the parametrisation strategy is shown in Figure 2.6.

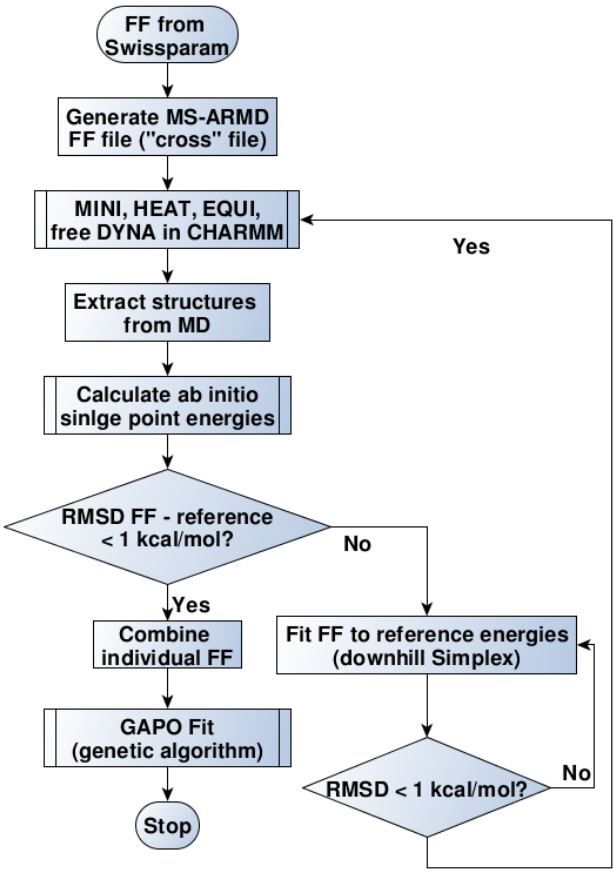


Figure 2.6.: Flowchart of the parametrisation strategy applied in this work. The starting point is an initial FF from Swissparam,[127] which is used for the construction of the MS-ARMD FF file, called "cross" file. This is used for minimization (MINI), heating (HEAT), equilibration (EQUI) and free dynamics (free DYNA) in the CHARMM program suit. From the free dynamics structures are extracted and their energies evaluated using a given *ab initio* reference method. If the FF yields an RMSD above 1 kcal/mol with respect to the reference method a downhill Simplex algorithm is used to fit the FF parameters to reproduce the reference energies. If the RMSD of the FF energies with respect to the reference energies drops below 1 kcal/mol the fitted FF is used to generate new structures using the same settings for the dynamics. Once the fitted FF yields an RMSD below 1 kcal/mol on the set of structures including the newly generated structures it is considered to be converged and combined with the remaining FFs to be passed to the GAPO fit GA that parametrises the GAPOs at the barrier region.

2.4. Classical Molecular Dynamics

Molecular dynamics are an extensively used method,[116,128] where the motion of atoms on the PES is calculated using Newton's equations of motion.[9,116] Those deterministic Newtonian mechanics use Newton's second law ($F = ma$; F: force, m: mass, a: acceleration) and the fundamental differential equation

$$F(r) = -\frac{\partial V}{\partial r} \tag{2.16}$$

which states that the force ($F(r)$) with respect to the position r is the derivative of the potential V. The position of each particle as a function of time (t) can be calculated using the following differential equation:

$$-\frac{\partial V}{\partial r} = m * \frac{\partial^2 r}{\partial t^2} \tag{2.17}$$

The acceleration is therefore defined as the second derivative of the position. The basic assumption used in MD is the ergodic hypothesis which states that the time averaging of one trajectory is equivalent to an average of the ensemble.[105] The way by which the coordinates are evaluated as a function of time in a MD simulation is by (dynamical) integrators, such as the Verlet integrators.[128] In the Verlet algorithm the acceleration can be described by either the forces or the potential.[105]

$$a = \frac{F}{m} = -\frac{1}{m}\frac{\partial V}{\partial r} \tag{2.18}$$

Two commonly applied schemes of the Verlet integrator are Leapfrog Verlet and Velocity Verlet (see Figure 2.7).[9,128] The Leapfrog algorithm uses expansions, starting with positions r_i and small half time steps and derives the velocities as well as the the coordinates with respect to time:

$$\begin{aligned} r_{i+1} &= r_i + v_{i+1/2}\delta t \\ v_{i+1/2} &= v_{i-1/2} + a_i \delta t \end{aligned} \tag{2.19}$$

This Leapfrog algorithm comes with the disadvantage that the coordinates and velocities are updated at different times (half a time step apart). This can be overcome by application of the velocity Verlet algorithm.

$$\begin{aligned} r_{i+1} &= r_i + v_i \delta t + \frac{1}{2} a_i \delta t^2 \\ v_{i+1} &= v_i + \frac{1}{2}(a_i + a_{i+1}) \delta t \end{aligned} \tag{2.20}$$

Here the coordinates and velocities are updated at the same time (see Figure 2.7).

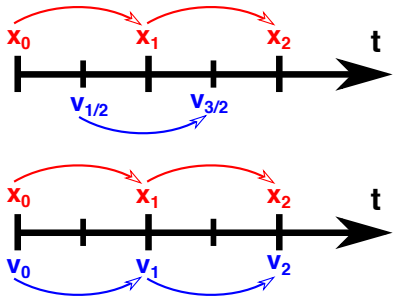


Figure 2.7.: Schematic representation of the coordinate and velocity updating of the Leapfrog Verlet (top) and the Velocity Verlet (bottom) algorithm.

MD simulations are performed in ensembles in which the number of atoms N as well as two other properties are kept constant. If additionally to the number of atoms the energy E and the volume V are kept constant, the resulting ensemble is called microcanonical (NVE). For the case where N, V and the temperature T are kept constant one achieves a canonical ensemble (NVT). In the isothermal-isobaric ensemble (NPT) the temperature and pressure P are kept constant, in addition to N.[9]

A general drawback of MD simulations is the requirement of extensive sampling in phase space (ergodicity hypothesis).

Umbrella Sampling

Conventional MD methods can sometimes infrequently sample regions of the conformational space by thermal fluctuations alone,[89] e.g. the high energy region at the TS.[105] However, if the reaction path is known (e.g from IRC calculations),[89] one can sample those regions by using, for example umbrella sampling (US). In US a harmonic biasing potential ($U(r)$) is added to the PES along a reaction coordinate.[89,105]

$$U(r) = k_u(r - r_0)^2 \tag{2.21}$$

where k_u is the force constant, r is the current value of the reaction coordinate and r_0 the predefined equilibrium value of the reaction coordinate for the given umbrella. Simulations performed with this added potential, assuming $U(r)$ is big, sample all degrees of freedom except for the reaction coordinate.[89] Through a number of simulations, analyzing all values of r_0 along the pre-defined reaction coordinate (see Figure 2.8), the total probability distribution function is obtained, the so called Potential of Mean Force (PMF).[89,105]

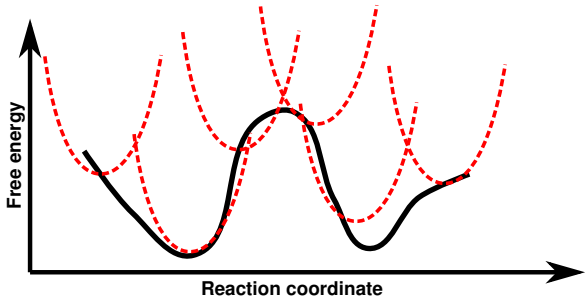
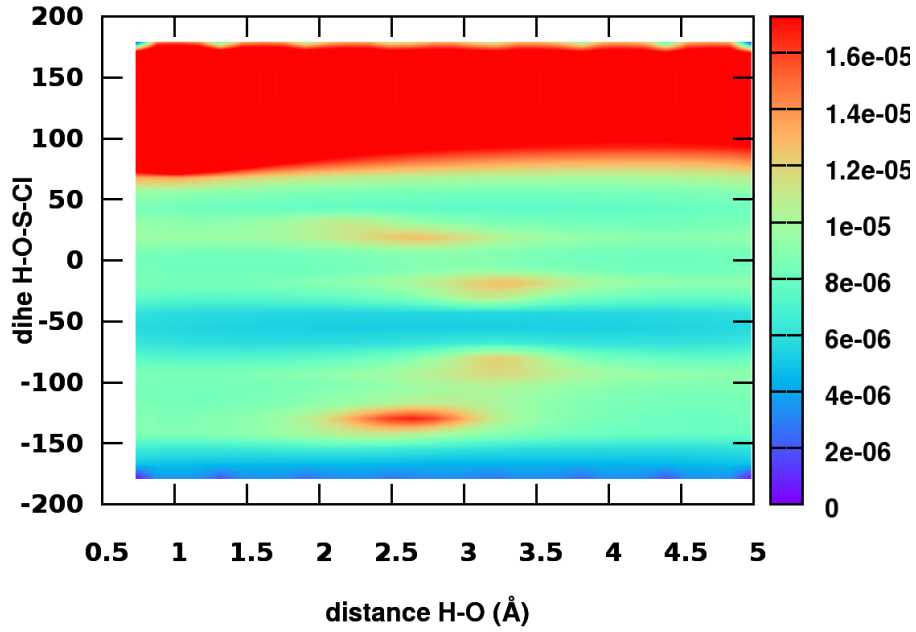


Figure 2.8.: Schematic representation of US simulations along a reaction coordinate. Each dashed line represents one umbrella with the same force constant k_u but different equilibrium values r_0 on the reaction coordinate.

PMFs are a useful tool for determining the free energy barrier of a reaction, which can be directly compared to values for the rate liming step obtained experimentally.

Part III.

Applications



Time discovers truth.

- *Lucius Annaeous Seneca* -

3. MS-ARMD for OH-Stretch Overtone Induced Photodissociation Dynamics

The absorption of infrared or visible light can pump energy into a vibrational overtone of the ground electronic state.[39,58] Of particular interest are those systems where an X-H (X = C, O, S, N) is excited.[58] In general the X-H stretch vibration, due to the large difference in mass between X and H, has a weak coupling to the others modes.[39,41] The interest in these vibrational overtone reactions is due to their importance in the atmosphere[65–68] but also because they are experimentally amenable.[58] If the intramolecular vibration relaxation (IVR) is slow in comparison to the unimolecular reaction, dissociation of the molecule can occur. This is the case for many systems in which the excited vibrational state is strongly coupled to the reaction coordinate.[58]

The dehydration reaction of H_2SO_4 plays an essential role in Earths atmosphere. The photofragmentation reaction plays an important role in the formation of atmospheric aerosols. Due to experimental difficulties for investigating the proposed OH-stretching overtone induced photodissociation,[129–131] HSO_3X[130,131] (X = F[26], Cl[27]) were proposed as proxy systems. MS-ARMD[15] simulations of HSO_3X were performed on a fully-dimensional PES. The PES, which describes H-transfer and HX-elimination, were fitted to MP2 calculations. The systems were set up similarly to the calculations of H_2SO_4, with changes to the maximum simulation time (increased to 2.5 ns), and on the number of independent trajectories (decreased to 5000). The OH-stretch excitation was achieved by instantaneously scaling the velocities of OH. Extensive simulation of OH-stretching overtone induced photodissociation of HSO_3X in the ns-time scale reveals dynamical effects on the dissociation reaction that differs for the two systems in question.

3.1. HSO_3Cl

In this work sulfurochloridic acid, a closely related molecule to H_2SO_4,[129,130] which is experimentally more liable was studied. The fragmentation energy is comparable to that of the dehydration reaction from H_2SO_4.[130] In order to validate previous finding and to better understand the dynamics of the fragmentation reaction MS-ARMD simulations were carried out. In order to provide a testable hypothesis for experiment the final states of the fragments were calculated.

Comparing the MS-ARMD results of the photofragmentation of HSO_3Cl to the ones of H_2SO_4,[25] a substantial difference in the percentage of reactive trajectories leading to elimination can be observed. For HSO_3Cl the yield of HCl elimination after 1 ns of simulation increases. Furthermore, the ratio between direct and indirect, i.e. dissociation without and with prior H-transfer, switches. For sulphuric acid the rate of intramolecular H-transfer is constant and secondary in comparison to direct water elimination. These differences can be explained by the transformation of symmetry and the more rapid elimination in H_2SO_4, respectively. The vibrational energy distribution of HCl, which is experimentally the most amenable observable, reveals a slight bimodal j-distribution for higher excitation degrees. This can be explained by differences in the redistribution of energy. For indirect elimination the excess energy goes more into other modes than for direct elimination. The latter therefore peaks at lower j values. For the simulations of HSO_3Cl it was further shown, by running simulations on an ensemble generated by Monte Carlo sampling, that the conclusions are reliable and independent of the choice of initial conditions.

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 6780

Received 27th November 2015,
Accepted 20th January 2016

DOI: 10.1039/c5cp07319g

www.rsc.org/pccp

HSO_3Cl: a prototype molecule for studying OH-stretching overtone induced photodissociation†

Juvenal Yosa Reyes,‡[ab] Sebastian Brickel,‡[a] Oliver T. Unke,[a] Tibor Nagy[ac] and Markus Meuwly*[a]

Vibrationally induced photodissociation in sulfurochloridic acid (HSO_3Cl) is found to be a viable process to form SO_3 and HCl from excitations of the OH-stretching overtone starting at ν_{OH} = 4. Reactive molecular dynamics simulations on a fully-dimensional potential energy surface fitted to MP2 calculations show that hydrogen transfer and HCl elimination compete with one another on the nanosecond time scale. Excitation with 5 and 6 quanta in the OH-stretch direct elimination of HCl is a dominant process on the several hundred picosecond time scale. At longer times, HCl formation is preceded by intramolecular hydrogen transfer and concomitant excitation of torsional degrees of freedom. As HSO_3Cl is a suitable proxy for H_2SO_4, which is relevant for weather and climate in the upper atmosphere, it is concluded that vibrationally induced photodissociation is a possible mechanism for H_2SO_4 decomposition. Final state energy distributions for different internal degrees of freedom are predicted which should be observable in laboratory measurements.

I. Introduction

The photochemistry of halogen- and sulfur-containing compounds which are involved in acid rain and aerosol formation in the troposphere and the stratosphere is of significance to the chemistry of the atmosphere. This led to extensive studies, both experimental and theoretical, concerning a wide range of chemical reactions relevant for climate and weather.[1] In 1982 Hofmann and Rosen suggested that the formation of condensation nuclei from sulfur compounds is dominated by photochemical processes and they related the periodic decrease and increase of these particles to the cycle of polar night and day around the poles.[2] This seasonal change in solar radiation affects directly the stratospheric aerosol layer by altering the rate of photochemical reactions of sulfur-containing compounds and thereby influences the climate and weather.

One of the central sulfur compounds in the stratospheric aerosol layer is sulfuric acid (H_2SO_4).[2–4] Sources for H_2SO_4 in the atmosphere are generally sulfur-containing compounds which are primarily emitted as dimethyl sulfide (DMS) and sulfur dioxide (SO_2).[5–9] These compounds are oxidized in the atmosphere to form sulfur trioxide (SO_3) and are subsequently hydrolyzed to H_2SO_4.[10,11] Among others, stratospheric sulfate aerosol has a global dimming effect in the atmosphere and increases the number of aerosol particles which can limit the effect and impact of global warming.[12] This makes H_2SO_4 a suitable candidate for climate engineering projects.[13] Additionally, the formation of a layer of dispersed aerosol found at high altitudes (between 15 to 35 km) which influences stratospheric ozone depletion, the so-called Junge layer,[14] can be explained by the cycling between SO_2 and H_2SO_4.[15–17]

Sulfur-containing compounds, in particular, the interplay between H_2SO_4 and SO_2, are important in the formation of atmospheric aerosols. The observed anomalous enhancement in particle concentration during springtime and the finding that the SO_2 concentration from the photolysis of H_2SO_4 is constant or increases with altitude[16] suggested that at high altitudes the photolysis of sulfuric acid is the source of SO_2.[18] However, experimental and computational investigations were unable to find the electronic absorption spectrum of H_2SO_4 down to 140 nm.[19–23] According to Lane and Kjaergaard[24] the lowest calculated electronic transition is at 139 nm. Hence, UV excitation from solar photons is improbable to produce SO_2 from H_2SO_4 and an alternative reaction mechanism for H_2SO_4 decomposition is required.

Experimentally and computationally it was previously found that vibrational excitation can induce bond-selective photodissociation.[25–27] It was therefore proposed that highly-excited OH-stretching states can lead to dissociation of H_2SO_4.[18] The vibrational spectrum of H_2SO_4 in the near-infrared and visible

[a] *Department of Chemistry, University of Basel, Klingelbergstrasse 80, Basel, Switzerland. E-mail: m.meuwly@unibas.ch*
[b] *Industrial and Systems Engineering Department, National University Bogotá, Carrera 45 # 26-85, Bogotá, Colombia*
[c] *IMEC, RCNS, Hungarian Academy of Sciences, Magyar tudósok körútja 2., Budapest, Hungary*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5cp07319g

‡ These authors contributed equally to this work.

regions has been intensely studied, both experimentally[19,28–31] and computationally.[22,23,31–33] Nevertheless it was not possible to experimentally verify that vibrational overtone-induced photodissociation is operative in H_2SO_4. This is due to particular experimental difficulties involving this species. By mixing known vapor flows of H_2O and SO_3 at elevated temperature (400–500 K) H_2SO_4 can be generated in the vapor phase ($H_2O + SO_3 \rightarrow H_2SO_4$).[19] However, H_2SO_4 has an extremely low vapor pressure, and exists in equilibrium with H_2O and SO_3 in the gas phase. Therefore this approach leads to difficulties in distinguishing the reactants and products.[29,34]

Because of the above mentioned difficulties and the widespread relevance of the atmospheric decomposition of sulfates, sulfuric acid derivatives have been considered to provide experimental evidence for characterizing OH-stretching induced dissociation in the gas phase.[29,34,35] Such candidate compounds are sulfurochloridic acid (HSO_3Cl) and sulfurofluoridic acid (HSO_3F) which are related to H_2SO_4,[29,34] but experimentally are more amenable. The dissociation energy to eliminate HCl and HF from HSO_3Cl and HSO_3F, respectively, is similar to that of H_2O from H_2SO_4 (32–40 kcal mol^{-1}).[34] Using a simple one-dimensional anharmonic oscillator local model, the OH overtone spectra of sulfuric acid derivatives have been computed, which found that excitation of the OH-stretching transition with $\Delta v \geq 4$ and $\Delta v \geq 5$, will provide sufficient energy for the photodissociation of HSO_3Cl and HSO_3F, respectively.[29,34,35] Finally, *ab initio* molecular dynamics (MD) simulations have demonstrated that the decomposition of these derivatives, similar to what has been proposed to H_2SO_4, is possible *via* excitation of higher OH-stretching overtone transitions.[29,34] At the semi-empirical PM3 level, it was found that for HSO_3Cl, excitation of $v \geq 4$, or $v \geq 5$ (PM3) can induce the elimination reaction.[34] The photodissociation of HSO_3F excitation of the $v \geq 7$ OH-stretching state is necessary with PM3, whereas excitation of the $v \geq 5$ overtone is sufficient at the MP2/TZP level.[34] The average reaction time for both HSO_3Cl and HSO_3F was of the order of 10 ps, determined from 32 trajectories.[34] However, previous reactive MD simulations on H_2SO_4 have demonstrated that for converged reaction time distributions several thousand trajectories are required.[32,33]

In order to characterize and understand the dynamics of vibrationally induced decomposition of HSO_3Cl and to guide future experimental efforts, Multi-Surface Adiabatic Reactive Molecular Dynamics (MS-ARMD) simulations were carried out.[36] This method is capable of treating multiple reaction pathways based on force fields to describe the adiabatic potential energy surface and the classical reaction dynamics of the system, and to run and analyze a statistically significant number of trajectories. Here, the dynamics on a fully dimensional reactive potential energy surface (PES) capable of following intramolecular H-transfer and HCl elimination in HSO_3Cl (see Fig. 1) is investigated. This provides the information to analyze the yields and determine rate coefficients along with the final state energy distribution in the products as a function of the vibrational overtone that was initially excited. This information can be directly compared with forthcoming experiments.

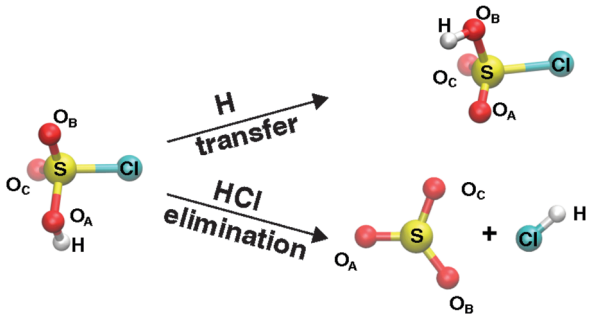


Fig. 1 Graphical representation of the two possible reaction pathways involving HSO_3Cl considered here. Top trace: intramolecular H-transfer ($O_AH \rightarrow O_BH$), bottom trace: HCl elimination.

II. Reactive molecular dynamics simulations

The global PES for the reactions of HSO_3Cl is based on reference electronic structure calculations at the MP2/6-311G++(2d,2p) level of theory. All *ab initio* calculations were performed using Gaussian09.[37] The energy at the equilibrium geometry of HSO_3Cl was chosen as the global zero of energy. The functional form of the force field which describes the HSO_3Cl, SO_3 and HCl states including the $SO_3\cdots HCl$ van der Waals complex has the following form:

$$\begin{aligned} V(x) = & \sum_{\text{bonds}} D_e\left(1-e^{-\beta(r-r_0)}\right)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 \\ & + \sum_{\text{Urey-Bradley}} K_{UB}(S-S_0)^2 \sum_{\text{dihedral}} H_\phi(1+\cos(n\phi-\delta)) \\ & + \sum_{\text{impropers}} K_\omega(\omega-\omega_0)^2 \\ & + \sum_{ij}\left\{\frac{n\varepsilon_{ij}}{m-n}\left[\left(\frac{r_{\min,ij}}{r_{ij}}\right)^m - \frac{m}{n}\left(\frac{r_{\min,ij}}{r_{ij}}\right)^n\right] + \frac{q_iq_j}{4\pi\varepsilon_0 r_{ij}}\right\} \end{aligned} \quad (1)$$

where r_0 is the equilibrium bond distance, D_e is the dissociation energy, and β controls the steepness and width of the Morse potential.[36] Additionally a generalized Lennard-Jones potential $\sum_{ij}\frac{n\varepsilon_{ij}}{m-n}\left[\left(\frac{r_{\min,ij}}{r_{ij}}\right)^m - \frac{m}{n}\left(\frac{r_{\min,ij}}{r_{ij}}\right)^n\right]$ is included, in order to better represent the van der Waals interaction between atoms in the reactive region.[36] The PESs of the three molecules (HSO_3Cl, SO_3 and HCl) together with the van der Waals complex ($SO_3\cdots HCl$) were parameterized independently. Specifically, the HSO_3Cl state is described by all terms in eqn (1), whereas for SO_3 only bonds, angles, Urey–Bradley, improper dihedrals and non-bonded terms, are needed. For HCl a bonded term and non-bonded terms (for the van der Waals complex) were required. For the reactive PES of the overall system the individual PESs were joined into a global reactive PES using the MS-ARMD switching method and were combined with Gaussian × polynomial (GAPOs) functions which provide the necessary flexibility to model the barrier crossing region.[36]

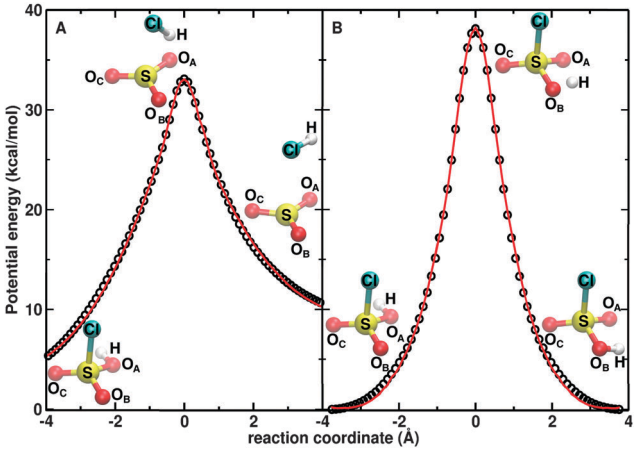


Fig. 2 Minimum energy path for (A) HCl elimination and (B) intramolecular H-transfer. Black circles represent the *ab initio* reference energies, red curves represent the FF energies including GAPOs.

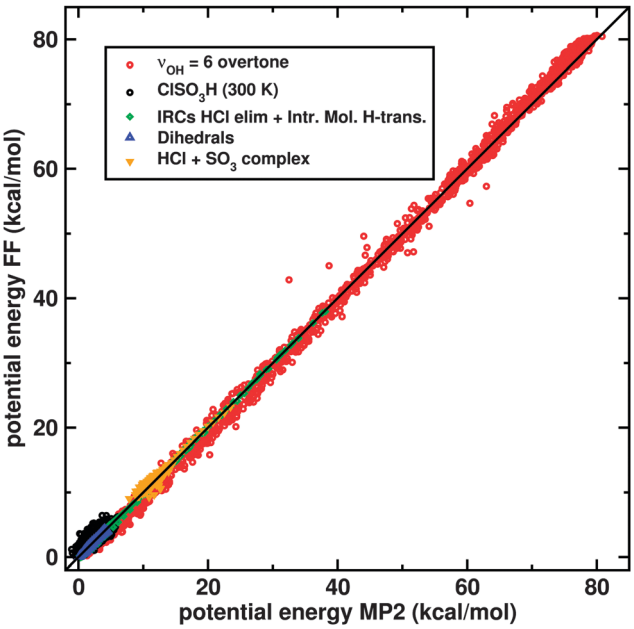


Fig. 3 Quality of the reactive PES. The reference data points are calculated at the MP2/6-311G(d,p) level of theory and compared to the parameterized surface.

The interpolation of the PESs is accomplished through energy-dependent weights $w_i(E)$.[36]

Details of the fitting are described in the ESI.† The optimization of the force field yields parameters for the global reactive PES which allows us to follow the H-transfer and HCl elimination reactions. The root-mean-square deviations (RMSD) between *ab initio* reference energies and FF energies are 0.22 and 0.29 kcal mol^{-1}, for the barrier region of HCl elimination and intramolecular H-transfer, respectively, see red curves in Fig. 2A and B.

The RMSD between the energy-minimized structures at the MP2/6-311G++(2d,2p) level and using the force field is 0.004 Å whereas for the $SO_3\cdots HCl$ van der Waals complex the RMSD is 0.053 Å. A comparison of equilibrium bond lengths, angles and dihedral angles is provided in Table SI (ESI†). The geometries in Table SI (ESI†) are also in good agreement with those computed using MP2/TZP (Dunning's triple-ζ+ polarization).[34] In addition, a normal mode analysis was performed in order to compare the experimentally observed vibrational frequencies with those from the MP2 and FF calculations, see Table SII (ESI†). The average differences between experimental,[38,39] MP2 and FF harmonic frequencies for HSO_3Cl are 4.56 and 4.71 cm^{-1}, respectively which establishes the quality of the force field.

Particularly relevant for the reaction rates are the barrier heights. The MP2/6-311G++(2d,2p) computed values are close to those determined previously using CCSD(T)/aug-cc-pV(T+d)Z.[34] For HCl elimination the value of 34.0 kcal mol^{-1} [34] compares with 33.1 kcal mol^{-1} from the present study. As the final global PES provides a good description of the equilibrium geometry, the harmonic frequencies and the barrier heights compared to results from *ab initio* calculations, the reactive PES is considered to be parameterized in a meaningful way. Contrary to an *ab initio* MP2/6-311G++(2d,2p) single point calculation, the reactive FF requires a fraction of the computational cost, which is highly relevant for dynamical studies.

The quality of the reactive PES is shown in Fig. 3. It provides a comparison between the energies of the reference MP2 calculations compared with the FF. For the representative structures from an equilibrium simulation at 300 K, the IRCs (HCl elimination and intramolecular H-transfer, obtained from individual scans) and the HCl + SO_3 van der Waals complex the absolute average difference between the reference MP2 energies and the FF is 0.52 kcal mol^{-1}.

All MD simulations were carried out using CHARMM[40] with provisions for bond-breaking and bond-formation through MS-ARMD.[36] Starting from a geometry optimized structure of HSO_3Cl, the system was heated to 300 K. The equations of motion were propagated using the leapfrog Verlet algorithm with a time step of Δt = 0.1 fs during 40 ps and equilibrated for 40 ps, followed by 250 ns of free dynamics simulations (*i.e.* constant energy *NVE*), where a structure was collected every 50 ps. The small time step is required to appropriately follow the rapid dynamics of the H-atom.

The non-equilibrium state was prepared by scaling the instantaneous velocity vector at the moment of excitation along the normal mode of interest which is the ν_{OH} mode in the present case. Similar to the situation in H_2SO_4 this mode has a predominantly local OH-stretching character. Such a scheme, which modifies only the kinetic energy, has already been successfully employed for H_2SO_4 and for studying proton transfer in small H-bonded complexes.[32,33,41] An alternative approach is to change both kinetic and potential energy by modifying positions and velocities.[42] While changing only the kinetic energy can lead to accelerated redistribution of vibrational energy, modification of positions and momenta requires a short equilibration period to avoid artifacts due to close proximity of atoms which may be a potential drawback of this approach.[42] Non-equilibrium excitation of a thermal ensemble as described above is closest to the situation encountered in a

laboratory experiment and was therefore preferred over generating quasiclassical (QCT) nonequilibrium initial conditions.

The OH stretch potential can be realistically described by a Morse potential for which the exact energy levels are known. The excitation energies for ν_{OH} = 3, 4, 5, and 6 are: 11 699 cm^{-1} (33.3 kcal mol^{-1}), 14 760 cm^{-1} (42.2 kcal mol^{-1}), 17 662 cm^{-1} (50.5 kcal mol^{-1}), and 20 405 cm^{-1} (58.3 kcal mol^{-1}) using D_e = 122.88 kcal mol^{-1} and β = 2.11 $Å^{-1}$ from the current O–H bond in the MS-ARMD potential. An alternative means to select initial conditions is described in the Results section.

III. Results

From 5000 independent trajectories with a maximum simulation time of 2.5 ns, more than 90% of the trajectories were reactive when ν_{OH} = 5 or 6 were excited (see Fig. 4). Both processes, (I) HCl elimination without prior H-transfer (dashed line), called direct, and (II) HCl elimination preceded by intramolecular H-transfer (dotted dashed line), called indirect, are observed. Direct HCl elimination is found for 0.2, 10.4, 28.5, and 42.8% of the cases for ν_{OH} = 3, 4, 5, and 6. HCl elimination preceded by intramolecular H-transfer was observed for 4.2, 62.0, and 54.9% of the cases for ν_{OH} = 4, 5 and 6 whereas for ν_{OH} = 3 this reaction channel was not observed. Intramolecular H-transfer without HCl elimination within 2.5 ns (leading to complete intramolecular vibrational energy redistribution (IVR), which strongly decreases the decay probability) was found for 10.2, 35.3, 7.3, and 0.08% of the cases for ν_{OH} = 3, 4, 5, and 6.

Excitation with ν_{OH} = 5 and 6 decomposition occurs on the hundred picosecond time scale. Overall 50% of the trajectories show HCl elimination within 500 ps for excitation with ν_{OH} = 6. For ν_{OH} = 5 the proportion is around 34% after 500 ps. The majority of these trajectories that show decomposition for ν_{OH} = 5 and 6 lead to direct HCl elimination, without previous H-transfer. On a longer time scale indirect elimination becomes dominant, as some time is required for "hydrogen hopping". For ν_{OH} = 3 and 4 decomposition of activated HSO_3Cl continues and much longer simulation times, particularly for the excitation with 3 quanta, would be required until all potential elimination reactions have occurred. It is worthwhile to mention that the mean collision time was estimated to be around 130 ns.[33] However, multiple 1000 simulations on this time scale become computationally prohibitive even with an empirical force field. The effect of deuteration was assessed by running 5000 trajectories for DSO_3Cl with ν_{OD} = 5 which corresponds to 13 546 cm^{-1} (36.2 kcal mol^{-1}). Within 2.5 ns, 2% of these trajectories showed DCl elimination, nearly all of them following process I. Hence, for the deuterated species much higher vibrational excitations are required for photodissociation which leads to a smaller cross section due to reduced overlap of the initial and final vibrational states than for the hydrogenated species.

Compared with a previous study on H_2SO_4[33] considerable differences in the number of reactive events for excitation with 5 or 6 quanta are found. For HSO_3Cl the percentage of trajectories which lead to HCl elimination (yield) is ≈60% for ν_{OH} = 5 after 1 ns, which is markedly increased compared to 25% in H_2SO_4 for the same OH stretching overtone within 1 ns, even though the barrier in H_2SO_4 is slightly lower. For sulfuric acid direct elimination was the dominant process for all excitation levels. Only 3.4% and 12.2% of the trajectories in H_2SO_4 for 5 and 6 quanta, respectively, were found to proceed *via* indirect elimination.[33] This can be explained by the fact that in H_2SO_4 direct elimination occurs more rapidly than in HSO_3Cl. As an example, for ν_{OH} = 5 the number of direct eliminations levels

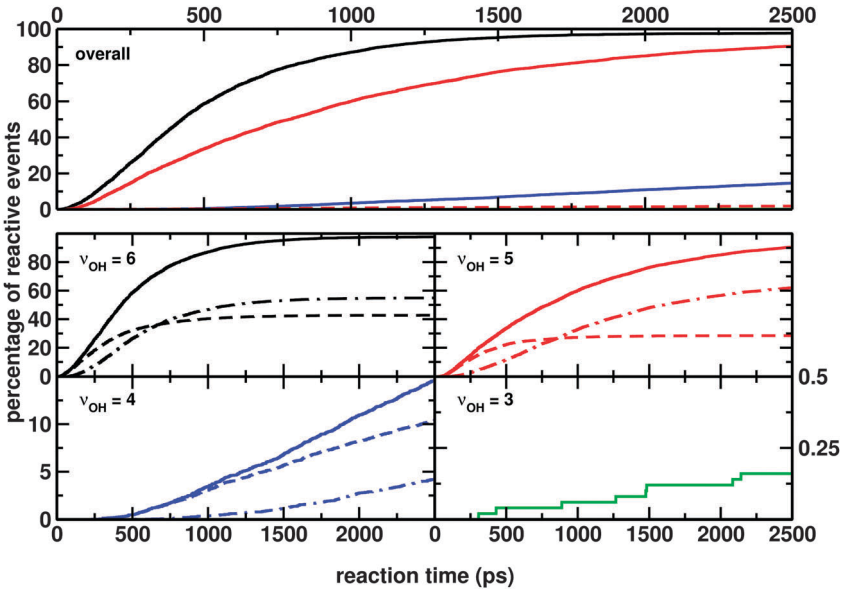


Fig. 4 Percentage of HCl elimination as a function of simulation time. Top panel: total number of reactive events (direct HCl elimination + HCl elimination preceded by intramolecular H-transfer). Green, blue, red and black curves correspond to excitation of ν_{OH} = 3, 4, 5 and 6, respectively. The red dashed curve corresponds to DCl elimination in DSO_3Cl with ν_{OD} = 5. Bottom panels: total number of reactive events (solid), direct HCl elimination (dashed), and HCl elimination preceded by intramolecular H-transfer (dotted dashed) for ν_{OH} = 3, 4, 5 and 6 in the same color code as in the top panel.

out (*i.e.* all eliminations on longer time scales are preceded by intramolecular H-transfer) after 300 ps in H_2SO_4 compared to 500 ps in HSO_3Cl.

The higher product yield for the same time after excitation (see previous paragraph) can be explained, in part, by the fact that the excitation energy in HSO_3Cl is higher than in H_2SO_4, which speeds up the elimination reaction: for ν_{OH} = 5 the difference between H_2SO_4 (47.2 kcal mol^{-1}) and HSO_3Cl (50.5 kcal mol^{-1}) is 3.3 kcal mol^{-1} whereas the barrier heights for H_2O and HCl elimination only differ by 1.1 kcal mol^{-1}, respectively. Furthermore, in H_2SO_4 two hindered rotors (HOSO) need to be in a relative position allowing for H_2O elimination, whereas for HCl elimination in HSO_3Cl only one such dihedral exists. On the other hand, the number of ways in which H_2O elimination can occur is larger (two –OH groups) for H_2SO_4 compared to HCl elimination in HSO_3Cl. Hence both energetic and entropic factors contribute to the differences in the product yields for the two systems. However, a detailed comparative analysis of these effects is outside the scope of the present work.

Most relevant for future experiments is the final state analysis of the products, which is also an observable.[43] In order to characterize the energy distribution after vibrational excitation and subsequent HCl elimination, the translational, rotational and vibrational energies in the reaction products were determined following established methodologies, as described in the ESI.† The analysis was carried out for all trajectories that experienced HCl elimination within 2.5 ns for ν_{OH} = 4, 5, and 6. The analysis for ν_{OH} = 3 was excluded due to the small number of reactive events. Final state analysis starts after HCl elimination when the sulfur–chlorine distance exceeds 20 Å. Trajectories were then continued for an additional 5 ps during which 500 snapshots for each trajectory were collected and analyzed. Translational, rotational, and vibrational energies as well as the total angular momentum were computed for every snapshot. Rotational and vibrational energies were averaged over the 500 snapshots. From these averages the probability distributions of the averaged values were determined.

Fig. 5 shows the distribution of translational, rotational, and vibrational energies of the products for ν_{OH} = 4 (blue), 5 (red), and 6 (black). The peak position of translational energies differs by less than 3 kcal mol^{-1} for the three excitation levels. For SO_3 a maximum peak for excitation with all quanta is found at about 5 kcal mol^{-1}. Rotational energy distributions for HCl peaks at about 3.5 kcal mol^{-1} (corresponding to a maximum angular momentum of j = 11), with a difference of the peak position of less than 1 kcal mol^{-1}. For SO_3 the rotational energy shifts towards higher energies for lower ν_{OH}. The vibrational energies show a similar behavior for HCl and SO_3. In both cases the vibrational energy distribution increases for higher overtone excitation. For SO_3 the maximum peaks at around 11, 15, and 19 kcal mol^{-1} for ν_{OH} = 4, 5 and 6, respectively.

The j-distribution (Fig. SIII (ESI†)) for SO_3 exhibits a single maximum around j = 60 to 70 for all investigated excitations. Contrary to that, for HCl a slight bimodal j-distribution for higher vibrational excitations, specifically for ν_{OH} = 6 is found. The two peaks can be assigned to the direct and indirect product channels. Direct elimination, with a peak around j = 10, is associated with higher vibrational energy, whereas indirect elimination leads to a higher j-value (around j = 12).

An alternative method to characterize final states from MD simulations, in particular, for vibrational degrees of freedom, is the Gaussian binning procedure.[44] This was applied to the diatomic product as this would also be most relevant from an experimental perspective. Gaussian binning involves assigning larger weights to final states with fractional vibrational quantum numbers x_i closer to an integer value n_i. For the present case of a diatomic and following ref. 44 the weight is $w = \sqrt{\beta}e^{-\beta(x-n)^2}$

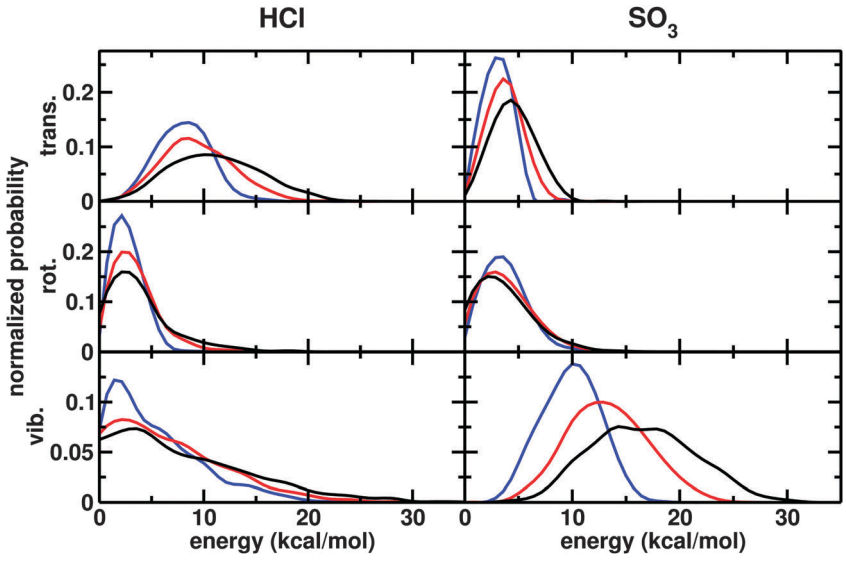


Fig. 5 Final state analysis for HCl and SO_3. Blue, red, and black curves correspond to excitation of ν_{OH} = 4, 5 and 6, respectively. The abbreviations trans., rot. and vib. stand for the translational, rotational, and vibrational energy, respectively.

where x and n are the fractional and integer quantum numbers of HCl in the product state, respectively. Parameter β should be chosen sufficiently large in order to reduce weights of final states near half-integer values and was $\beta = 20$ in the present case. For excitation of HSO_3Cl with $\nu = 4$ the final state population for the HCl stretching mode is 43% ($v_{HCl} = 0$), 30% ($v_{HCl} = 1$), 18% ($v_{HCl} = 2$), and 9% ($v_{HCl} \geq 3$). This only differs marginally from weighting all final states equally ($w = 1$) and supports the final state analysis discussed above.

In order to validate the results, an alternative method to generate initial conditions was employed. Instead of starting from an equilibrium ensemble at 300 K, energies were drawn from a Maxwell–Boltzmann distribution using Metropolis Monte Carlo (MC) sampling and initial conditions generated according to an algorithm developed by Schranz *et al.*[45] For each drawn energy E, spatial coordinates are selected from the final state of a Markov chain generated by a Metropolis-type procedure.[46] Then, random momentum vectors are generated from a Gaussian distribution for each atom. Finally, the momentum vectors are rescaled such that the total kinetic energy $T = E - V$, where V is the potential energy of the generated configuration. For convenience, all generated conditions were transformed to a center-of-momentum frame. The potential energy distributions for the two different ensembles are similar to one another. The distribution of the potential energy for the ensembles generated from the MC sampling shows a maximum peak slightly closer to $\frac{15}{2}k_B T$, which is the thermodynamic value.

Fig. 6 shows the decomposition probability for 1000 trajectories generated using the MC method. The excitation of the OH bond was applied in the same way as for the previous ensemble. Calculations with $\nu_{OH} = 3$ were excluded due to the small number of reactions found for the original set of initial conditions (see Fig. 4). Qualitatively, the reaction rates are comparable to those from simulations in which excitation of the equilibrium ensemble is used, see Fig. 4. Also, the relative importance of direct HCl elimination and elimination preceded by H-transfer to neighboring oxygen atoms is comparable. Quantitatively, the decomposition probability for all excitation is reduced by 25% to 30% compared to simulations using initial conditions taken from *NVE* simulations. Furthermore, the reaction slows down, as can be seen by comparing, for example, the point at which excitation with $\nu_{OH} = 6$ reaches 50% reactivity. For the results shown in Fig. 6 using the alternative set of initial conditions, this occurs at around 1 ns compared to $\approx$500 ps for the original choice of initial conditions, see Fig. 4. The same is found for the crossing point between direct and indirect elimination with $\nu_{OH} = 5$ and 6 which is shifted towards longer times.

The trajectories run under the Monte Carlo initial conditions were analyzed in the same fashion as the original 5000 trajectories, see Fig. SIV (ESI†). The energy distributions after HCl elimination are slightly shifted towards higher energies when compared to Fig. 5. However, the distributions remain qualitatively similar. As for the vibrationally induced decomposition reaction, conclusions drawn from the final state analysis are robust with respect to the two very different procedures to generate initial conditions for the non-equilibrium simulations. For excitation with $\nu_{OH} = 4$ beating patterns are observed for both types of initial conditions (but particularly pronounced with those from Monte Carlo sampling). However, interpretation of those would be premature, also because the number of reactive trajectories for this level of excitation is only around 15%.

The decomposition reactions preceded by intramolecular H-transfer differ from direct HCl elimination primarily by

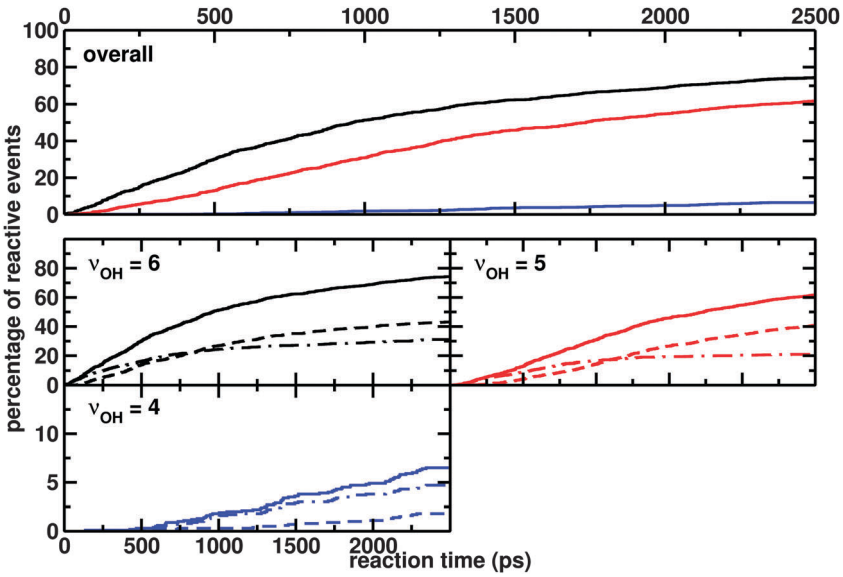


Fig. 6 Reaction kinetics from analyzing 1000 trajectories with initial conditions from the Monte Carlo sampling. (top panel) Total number of reactive events (direct HCl elimination + HCl elimination preceded by intramolecular H-transfer). Blue, red and black curves correspond to excitation of $\nu_{OH} = 4$, 5 and 6, respectively. (bottom panels) Total number of reactive events (solid), direct HCl elimination (dashed), and HCl elimination preceded by intramolecular H-transfer (dotted dashed) for $\nu_{OH} = 4$, 5 and 6 in the same color code as in the top panel.

the time scale on which dissociation occurs, see Fig. 4 and 6. In order to characterize the differences between these two reaction pathways, geometrical coordinates were sought which can be used to further analyze indirect elimination. Visual inspection of the dynamics suggested that the H–Cl–S–O_A dihedral angle ϕ could be a meaningful choice. The dihedral angle ϕ can distinguish between the three oxygen atoms in a symmetric way (with O_A on the mirror axis). The distance O_A helps us to further distinguish the reactive trajectories, especially with respect to the position of the hydrogen atom to its acceptor. The trajectories can be grouped into those in which during the dynamics the hydrogen atom binds to oxygen atoms O_A and O_B, O_A and O_C, or to all three oxygen atoms.

Probability distributions $p(r_{HO_A},\phi)$ from analyzing 1000 trajectories, taken from the ensemble in which excitation of the equilibrium was used, for each of the excitation levels are reported in Fig. 7. Each column displays, from top to bottom, H-transfer to (I) both oxygen atoms, (II) O_B only, (III) O_C only before HCl elimination occurs. The columns report results for the increasing number of OH-quanta from left to right. Each panel was individually normalized. For excitation with ν_{OH} = 4 (left column) only a small number of trajectories show indirect HCl elimination: for cases (I) to (III) there are 2, 39 and 10 events, respectively. For excitation with 5 quanta these numbers increase to 295, 525, and 104 trajectories whereas for excitation with ν_{OH} = 6 cases (I) to (III) occur for 356, 358, and 86 runs. Hence, for excitations with 4 to 6 quanta 90%, 10% and 20% of the trajectories, respectively, lead to HCl elimination without transfer between neighboring oxygen atoms.

The distribution of angle ϕ reveals two states for each possible H–O_X (X = A, B, C) bond, see red crosses with corresponding molecular models in Fig. 7. These states correspond to two synperiplanar positions of the hydrogen atom with respect to O_X. The probability distributions for the two states at $\phi = \pm 25^\circ$ and short HO_A bond length consist of the positions of the hydrogen atom sampled (a) in the equilibrium trajectories of HO_ASO_2Cl before excitation and (b) during the migratory dynamics after vibrational excitation. This explains the high probabilities in these two positions. The distribution of the H–O_A bond length broadens for higher vibrational excitations as the vibrational energy along the OH-bond increases (see left sides in the panels in Fig. 7). With ν_{OH} = 4 the hydrogen remains longer in a specific orientation since the HO_AS rotation barrier is rarely overcome. This – and the small number of trajectories in this class – leads to the asymmetric distribution along ϕ in the top left corner. For excitation with ν_{OH} = 5 the distribution $p(r_{HO_A},\phi)$ for trajectories sampling all oxygen atoms is most symmetrical. This indicates that the dihedral

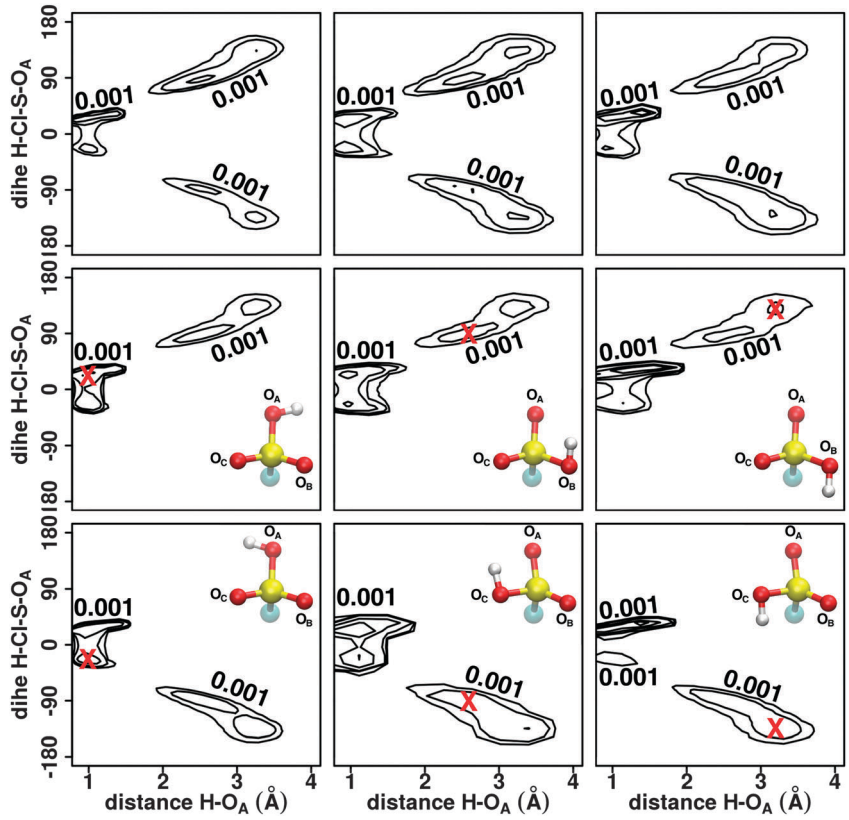


Fig. 7 Probability distribution $p(r_{HO_A},\phi)$, with ϕ as the dihedral angle between H–Cl–S–O_A, shown as contour lines. From left to right excitation with ν_{OH} = 4, 5, and 6. From top to bottom indirect trajectories showing H-transfer to both oxygen atoms, to O_B only, and to O_C only. The molecular structures represent the geometry of the region marked with a red cross (for each panel individually). $p(r_{HO_A},\phi)$ was normalized to unity for each panel and the lowest isocontour is labelled in each case.

barrier can be overcome and that extensive sampling of all oxygen positions by the hydrogen atom is possible before elimination. Finally, with ν_{OH} = 6 the distribution in the top right corner is again less symmetric as the residence time of the hydrogen on each oxygen before elimination is shortened due to the high excitation energy. The middle column of Fig. 7 illustrates the increasing H–O_A amplitude as the number of quanta is increased from left to right. From these probability distributions it is evident that coupling between the OH-stretching vibration and the torsional degrees of freedom is an essential pathway for energy exchange which influences the rates and yields of HCl elimination on the picosecond time scale.

IV. Conclusion

In summary, the present work establishes that vibrational excitation of OH-overtones with 3 to 6 quanta leads to HCl elimination on the sub-nanosecond to ns time scale for HSO_3Cl. For short times (a few hundred ps) after vibrational excitation this process is usually direct, *i.e.* without previous intramolecular H-transfer and leads to characteristic energy distributions in the diatomic fragment which should be experimentally detectable. Direct and indirect HCl elimination occurs whereby the latter is driven by appreciable excitation of torsional degrees of freedom and dominates at longer times. It is shown that the conclusions from the simulations are robust towards the particular choice of initial conditions. Experimentally corroborating such a reaction mechanism will be relevant for atmospheric chemistry at large, since under the conditions in the stratopause around the poles vibrationally excited states can survive sufficiently long for decomposition to occur. The reaction is fast (sub-ns to ns time scale) compared to collisionally induced quenching, which occurs on the hundred ns time scale.[18,33] Hence, there is only little competition for the dissociative channel. The present work provides testable hypotheses for experiments and establishes a vibrationally driven photodissociation mechanism for HSO_3Cl which is a meaningful proxy for H_2SO_4 relevant in atmospheric chemistry.

Acknowledgements

This work was supported by the Swiss National Science Foundation through grants 200021-117810, the NCCR MUST (to MM), an EU-COFUND grant (to TN), and the University of Basel. JY is grateful to the Patrimonio Autnomo Fondo Nacional Francisco José de Caldas and National University for financial support.

References

1 B. J. Finlayson-Pitts, *Anal. Chem.*, 2010, **82**, 770–776.
2 D. J. Hofmann and J. M. Rosen, *Nature*, 1982, **297**, 120–124.
3 F. Raes, R. V. Dingenen, E. Vignati, J. Wilson, J.-P. Pataud, J. H. Seinfeld and P. Adams, *Atmos. Environ.*, 2000, **34**, 4215–4240.
4 U. Pöschl, *Angew. Chem., Int. Ed.*, 2005, **44**, 7520–7540.
5 R. Charlson, J. Lovelock, M. Andreae and S. Warren, *Nature*, 1987, **326**, 655–661.
6 R. Kiene, *Nature*, 1999, **402**, 363–368.
7 R. Simó and C. Pedrós-Alió, *Nature*, 1999, **402**, 396–399.
8 B. Huebert, *Nature*, 1999, **400**, 713–714.
9 K. Capaldo, J. Corbett, P. Kasibhatla, P. Fischbeck and S. Pandis, *Nature*, 1999, **400**, 743–746.
10 L. J. Larson, M. Kuno and F. M. Tao, *J. Chem. Phys.*, 2000, **112**, 8830–8838.
11 F. C. Charles and J. F. Ian, *Atmos. Environ.*, 1999, **33**, 1352–2310.
12 P. Rasch, S. Tilmes, R. Turco, A. Robock, L. Oman, C.-C. Chen, G. Stenchikov and R. Garcia, *Philos. Trans. R. Soc. London, Ser. A*, 2008, **306**, 4007–4037.
13 K. Caldeira and L. Wood, *Philos. Trans. R. Soc. London, Ser. A*, 2008, **366**, 4039–4056.
14 D. J. Hofmann and J. M. Rosen, *Geophys. Res. Lett.*, 1985, **12**, 13–16.
15 J. Zhao, O. B. Toon and R. P. Turco, *J. Geophys. Res.: Atmos.*, 1995, **100**, 5215–5227.
16 C. P. Rinsland, M. R. Gunson, M. K. W. Ko, D. W. Weisenstein, R. Zander, M. C. Abrams, A. Goldman, N. D. Sze and G. K. Yue, *Geophys. Res. Lett.*, 1995, **22**, 1109–1112.
17 M. J. Mills, O. B. Toon and S. Solomon, *Geophys. Res. Lett.*, 1999, **26**, 1133–1136.
18 V. Vaida, H. G. Kjaergaard, P. E. Hintze and D. J. Donaldson, *Science*, 2003, **299**, 1566–1568.
19 J. B. Burkholder, M. Mills and S. McKeen, *Geophys. Res. Lett.*, 2000, **27**, 2493–2496.
20 P. E. Hintze, H. G. Kjaergaard, V. Vaida and J. B. Burkholder, *J. Chem. Phys.*, 2003, **107**, 1112–1118.
21 S. J. Wrenn, L. J. Butler, G. A. Rowland, C. J. Knox and L. F. Phillips, *J. Chem. Phys.*, 1999, **129**, 1112–1118.
22 H. G. Kjaergaard, J. R. Lane, A. L. Garden, D. P. Schofield, T. W. Robinson and M. Mills, *Adv. Quantum Chem.*, 2008, **55**, 1112–1118.
23 T. W. Robinson, D. P. Schofield and H. G. Kjaergaard, *J. Chem. Phys.*, 2003, **118**, 7226–7232.
24 J. R. Lane and H. G. Kjaergaard, *J. Phys. Chem. A*, 2008, **112**, 4958–4964.
25 F. F. Crim, *Annu. Rev. Phys. Chem.*, 1984, **35**, 657–691.
26 T. Uzer, J. Hynes and W. Reinhardt, *J. Chem. Phys.*, 1986, **85**, 5791–5804.
27 F. F. Crim, *J. Phys. Chem.*, 1996, **100**, 12725–12734.
28 D. K. Havey, K. J. Feierabend and V. Vaida, *THEOCHEM*, 2004, **680**, 243–247.
29 J. R. Lane, H. G. Kjaergaard, K. L. Plath and V. Vaida, *J. Phys. Chem. A*, 2007, **111**, 5434–5440.
30 P. E. Hintze, K. J. Feierabend, D. K. Havey and V. Vaida, *Spectrochim. Acta, Part A*, 2005, **61**, 559–566.
31 K. J. Feierabend, D. K. Havey, S. S. Brown and V. Vaida, *Chem. Phys. Lett.*, 2006, **420**, 438–442.
32 J. Yosa and M. Meuwly, *J. Phys. Chem. A*, 2011, **115**, 14350–14360.
33 J. Yosa, T. Nagy and M. Meuwly, *Phys. Chem. Chem. Phys.*, 2014, **16**, 18533–18544.

34 P. Gupta, J. R. Lane and H. G. Kjaergaard, *Phys. Chem. Chem. Phys.*, 2010, **12**, 8277–8284.
35 J. R. Lane and H. G. Kjaergaard, *J. Phys. Chem. A*, 2007, **111**, 9707–9713.
36 T. Nagy, J. Yosa and M. Meuwly, *J. Chem. Theory Comput.*, 2014, **10**, 1366–1375.
37 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision D.01*, Gaussian Inc., Wallingford, CT, 2009.
38 S. M. Chackalackal and F. E. Stafford, *J. Am. Chem. Soc.*, 1966, **88**, 4815–4819.
39 T. Shimanouchi, *Tables of molecular vibrational frequencies. Consolidated volume II*, 1977, http://link.aip.org/link/JPCRBU/v6/i3/p993/s1{\&}Agg=doi.
40 B. R. Brooks, C. L. Brooks, III, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545–1614.
41 M. Meuwly, A. Müller and S. Leutwyler, *Phys. Chem. Chem. Phys.*, 2003, **5**, 2663–2672.
42 P. N. Nguyen and G. Stock, *J. Chem. Phys.*, 2003, **119**, 11350–11358.
43 M. P. Grubb, M. L. Warter, H. Xiao, S. Maeda, K. Morokuma and S. W. North, *Science*, 2012, **335**, 1075–1079.
44 L. Bonnet and J. Rayez, *Chem. Phys. Lett.*, 1997, **277**, 183–190.
45 H. Schranz, S. Nordholm and G. Nyman, *J. Chem. Phys.*, 1991, **94**, 1487–1498.
46 N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087–1092.

Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

HSO_3Cl: A Prototype Molecule for Studying OH-stretching Overtone Induced Photodissociation

Juvenal Yosa Reyes[1,2,#], Sebastian Brickel [1,#], Oliver Unke[1], Tibor Nagy[1,3], Markus Meuwly[1,*]

[1]Department of Chemistry, University of Basel, Klingelbergstrasse 80, Basel, Switzerland

[2]Industrial and Systems Engineering Department, National University Bogotá, Carrera 45 # 26-85, Bogotá, Colombia

[3]IMEC, RCNS, Hungarian Academy of Sciences, Magyar tudósok körútja 2., Budapest, Hungary

[#]These authors contributed equally to this work

(Dated: January 18, 2016)

Abstract

Initial parameters for bonds (S=O, S−O and O−H), angles (O=S=O, O−S=O and S−O−H), dihedral (O=S−O−H) and Urey-Bradley (O=S=O, O−S=O and S−O−H) for HSO_3Cl were obtained from previous studies on H_2SO_4.[1]

MS-ARMD employs generalized-exponent Lennard-Jones potentials (or Mie potentials)[2] which are used for HSO_3Cl, between 1-4 (H−Cl and H−O_B or O_C) neighboring atoms.[3] Point charges were obtained by fitting the molecular electrostatic potential (ESP) using the CHELPG scheme within Gaussian09.[4] Initial values for all other parameters were taken or calculated (i.e. for generalized Lennard-Jones) from SWISSPARAM.[5]

Starting with these initial parameters, representative geometries were collected for each molecule (HSO_3Cl, HCl and SO_3) from MD simulations, which cover the regions of the multidimensional PES relevant to the equilibrium dynamics (in total of 4346). A MD simulation was started from the optimized geometry, heated to 300 K, and subsequently equilibrated for 40 ps, followed by 10 ns of free dynamics at constant total energy. To cover the regions of the phase space where the dynamics are sampled after excitation, the above protocol was followed after heating up to 1000 K. Starting along the three trajectories, approximately 1000, 1004 and 2041 (1000 at 300 K and 1041 at 1000 K) geometries were selected for HCl, SO_3 and HSO_3Cl, respectively. Additionally, a rigid scan along the dihedral angles Cl-S-O_A-H in HSO_3Cl was performed resulting in an extra 36 reference data points.

For every representative structure a single point energy was computed at the MP2/6-311G+(2d,2p) level. The energies and structures were subsequently used in a fitting procedure to a functional form:

$$\begin{aligned} V(\mathbf{x}) = & \sum_{\text{bonds}} D_e(1-e^{-\beta(r-r_0)})^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 \\ & + \sum_{\text{Urey-Bradley}} K_{\text{UB}}(S-S_0)^2 \sum_{\text{dihedral}} H_\phi(1+\cos(n\phi-\delta)) \\ & + \sum_{\text{impropers}} K_\omega(\omega-\omega_0)^2 \\ & + \sum_{ij}\left\{\frac{n\varepsilon_{ij}}{m-n}\left[\left(\frac{r_{\min,ij}}{r_{ij}}\right)^m - \frac{m}{n}\left(\frac{r_{\min,ij}}{r_{ij}}\right)^n\right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}\right\} \end{aligned} \quad (1)$$

in order to optimize the FF parameters. For this purpose, a downhill SIMPLEX algorithm[6] was

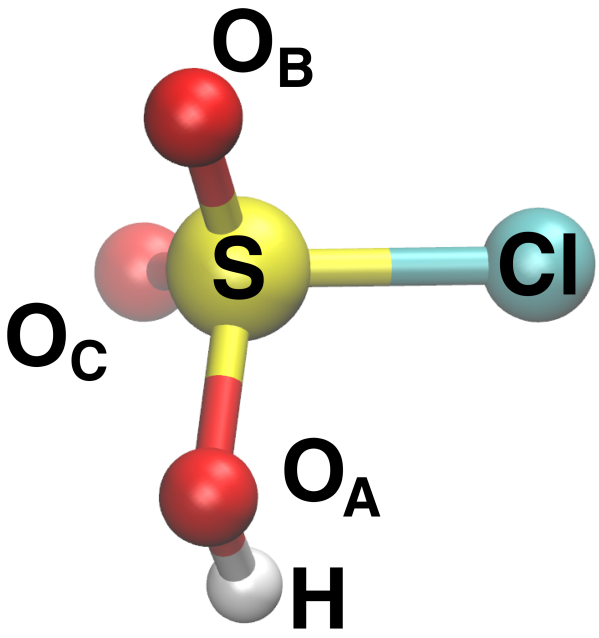


FIG. S-I: Structure of HSO_3Cl with labels.

used.

Similar to H_2SO_4, HSO_3Cl can follow two different pathways: intramolecular H-transfer and HCl elimination ($HSO_3Cl \rightarrow HCl + SO_3$), see Figure 1 in the manuscript.[1,7,8]

With the refined FF parameters for HCl and SO_3, the generalized Lennard-Jones (GLJ) parameters and atomic charges (last term in Eq. 1) were fitted in order to describe the intermolecular interactions for the $HCl\cdots SO_3$ van der Waals complex. Initial values for the GLJ parameters are obtained from the Lennard-Jones parameters between pairs of atoms i and j using the Lorentz-Berthelot combination rules in which ε_{ij} is the geometric mean of ε_i and ε_j and $r_{\mathrm{min},ij}$ is the arithmetic mean of $r_{\mathrm{min},i}$ and $r_{\mathrm{min},j}$. For the GLJ potential, ε_{ij} (> 0) and $r_{\mathrm{min},ij}$ are also the well-depth and the corresponding separation at this energy minimum, respectively.

With this initial parameter set, a new MD simulation for HSO_3Cl was run at 250 K to avoid dissociation. Another 1000 structures were collected and single point energies were computed at the MP2 level and included in the fit. The total number of structures that were used for the final fitting was 3376.

Following the same protocol as for H_2SO_4[1,3], surfaces for the reactant (HSO_3Cl) and products ($HCl + SO_3$) need to be combined and a permutation invariant force fields for the different, but chemically equivalent, product channels are needed for the global surface. In total three different surfaces characterize the intramolecular H-transfer states and one surface is used for HCl elimination. In MS-ARMD the FFs for the individual states i (connectivities) are mixed according to

$$\sum_{i=1}^{n} w_i(\vec{x}) V_i(\vec{x}) \tag{2}$$

where $w_i(\vec{x})$ are normalized weights of the raw weights $w_{i,0} = \exp\Big(- \big(V_j(\vec{x}) - V_{min}(\vec{x})\big)/\Delta V \Big)$, depending on the i-th PES $V_i(x)$. However, simple mixing will lead to discontinuous behavior in the barrier region[3] and typically overestimate the barrier, see features ($*$) in Figure S-II (lower panel). In order to address this point, Gaussian $\times$ polynomial functions (GAPOs) are used. Reference data in the barrier region and along the minimum energy paths (MEPs) were determined from quadratic synchronous transit (QST2) calculations at the MP2/6-311G+(2d,2p) level and the GAPOs were fitted to them while freezing the remaining FF terms at their previously optimised values. More detailed information about the GAPO function can be found in Nagy et al.[3] The barrier height of the intramolecular H-transfer is fitted with three GAPOs, each one with a polynomial of order four, leading to a symmetric intramolecular MEP. HCl elimination is represented with three GAPOs, each one with a polynomial of order five.

A comparison between the MP2 reference energies and the FF with and without GAPOs is provided in Figure S-II. Figure S-II (lower panel) shows the potential energy difference between the FF with and without the corrected GAPOs function. The main difference is shown in those zones where the two IRCs are located. Deviations less than 0.16 kcal/mol are found for the structures collected after excitation with $\nu_{OH} = 6$ (see Figure S-II), while for the thermal equilibrium at 300 K the effect of the GAPO functions is negligible. In that way one is sure that the correction functions only affects the energy of those structures close to the transition state.

Table S-I compares the geometric parameters for the equilibrated structures of the three molecules, obtained with MP2 and the FF. For the two photofragments there are no deviations between the *ab initio* and the FF bond lengths and angles. For HSO_3Cl slight but insignificant deviations were

found for bond length, angles, and dihedrals. Overall, the structures are well reproduced.

Table S-II shows the computed vibrational frequencies for HSO_3Cl and SO_3. All MP2 frequencies were scaled with a factor of 0.97 (MP2/6-31+G(d,p)[9] and MP2-fc/6-311G(d,p)[10]). As the FF is based on MP2 calculations, the vibrational frequencies from the normal mode analysis using the FF were also scaled with the same scaling factor. Comparison between the two sets of values shows a non-systematic difference between the MP2 and the FF energies. These differences are small for both molecules. When comparing the values obtained with MP2 and FF with experimental values it was found that the MP2 is somewhat closer to the measured values than the FF.

Final State Analysis: For the final state analysis, the total energy of the system was decomposed into translational (E_{trans}), rotational (E_{rot}), and vibrational (E_{vib}) components of each fragment. The translational energy is $E_{\text{trans}} = \frac{1}{2}Mv_{\text{CM}}^2$ where M and v_{CM} denote the mass and the velocity of the center of mass for each fragment (SO_3 and HCl). Following classical mechanics, the angular momentum vector of a reaction product (with N atoms) is $\mathbf{L} = \sum_{i=1}^{N} m_i \left(\mathbf{r}_i - \mathbf{r}_{\text{CM}}\right) \times \left(\mathbf{v}_i - \mathbf{v}_{\text{CM}}\right)$. Here, $\mathbf{r}_{\text{CM}}$ and $\mathbf{v}_{\text{CM}}$ denote the position and the velocity vectors of the center of mass for each fragment. This yields the rotational energy $E_{\text{rot}} = \frac{1}{2}\mathbf{L}^{\text{T}}\Theta^{-1}\mathbf{L}$ where $\Theta = \sum_{i=1}^{N} m_i \left(\mathbf{E}r_i^2 - \mathbf{r}_i\mathbf{r}_i^{\text{T}}\right)$ is the moment of inertia tensor of the fragment and $\mathbf{E}$ is the unit matrix. HCl as a linear molecule can only rotate around axes through its center of mass which are orthogonal to the HCl bond.

The angular momentum quantum number j is calculated from the classical-quantum correspondence relationship $L^2 = j(j+1)\hbar^2$ in which $\hbar$ is Planck's constant. The orbital angular momentum ($\mathbf{L}_{\text{orbital}}$) can be determined from the individual angular momenta of the fragments and from the conservation of total angular momentum.

From E_{trans} and E_{rot} the vibrational energy of each product is determined through the relationship $E_{\text{tot}} = E_{\text{pot}} + E_{\text{kin}}$. Utilising that E_{vib} is a part of E_{kin}, relative to the center of mass with addition of the intramolecular potential energy of the given reaction product, the vibrational energy can be expressed as $E_{\text{vib}} = E_{\text{kin}} - E_{\text{trans}} - E_{\text{rot}} + V$. Once the two fragments have moved sufficiently far from the interaction region (here taken as 20 Å between the SO_3-sulfur and the HCl-chlorine), the translational and rovibrational energy of each fragment will be constant. As the vibrational and rotational modes in each fragment keep on exchanging energy if the angular momentum is

nonzero, these quantities are averaged over periods much longer (i.e. 5 ps) than the characteristic time for vibrations within the molecules.

TABLE S-I: Comparison of MP2 and force field geometric parameters for the equilibrium of HSO_3Cl, SO_3 (symmetry D_{3h}, X = A, B, C) and HCl.

	MP2	**Force Field**
	HSO_3Cl	
Bonds (Å)		
S-Cl	2.066	2.074
S-O_B	1.432	1.429
S-O_C	1.424	1.428
S-O_A	1.596	1.602
O_A-H	0.969	0.978
Angles (°)		
O_C-S-O_B	124.5	124.3
O_C-S-O_A	109.1	109.0
O_B-S-O_A	106.3	106.4
Cl-S-O_B	107.4	107.3
Cl-S-O_C	106.6	106.8
Cl-S-O_A	100.3	100.6
S-O_A-H	107.7	108.1
Dihedral (°)		
O_B-S-O_A-H	164.8	164.6
O_C-S-O_A-H	28.2	28.3
Cl-S-O_A-H	-83.6	-83.7

	MP2	**Force Field**
	SO_3	
Bonds (Å)		
S-O_X	1.44	1.44
Angles (°)		
O_X-S-O_Y	120.0	120.0
	HCl	
Bonds (Å)		
Cl-H	1.27	1.27

TABLE S-II: computed and experimental vibrational frequencies for HSO_3Cl and SO_3. MP2 and FF scaled by a factor of 0.97.

MP2 cm^{-1}	**Experimental cm^{-1}**	**FF cm^{-1}**
HSO_3Cl		
252.5		250.7
277.7		306.8
355.7		315.6
385.4		376.6
444.3		493.9
479.3	513.0	507.0
578.4	614.0	538.8
798.3	852.0	799.1
1149.1		1106.8
1170.0	1235.0	1141.0
1411.5	1455.0	1364.5
3663.9	3587.0	3656.5
SO_3		
474.4	498	457.2
509.6 509.8	530	496.7 496.7
1020.6	1065	1080.3
1366.4 1367.3	1391	1363.1 1363.2

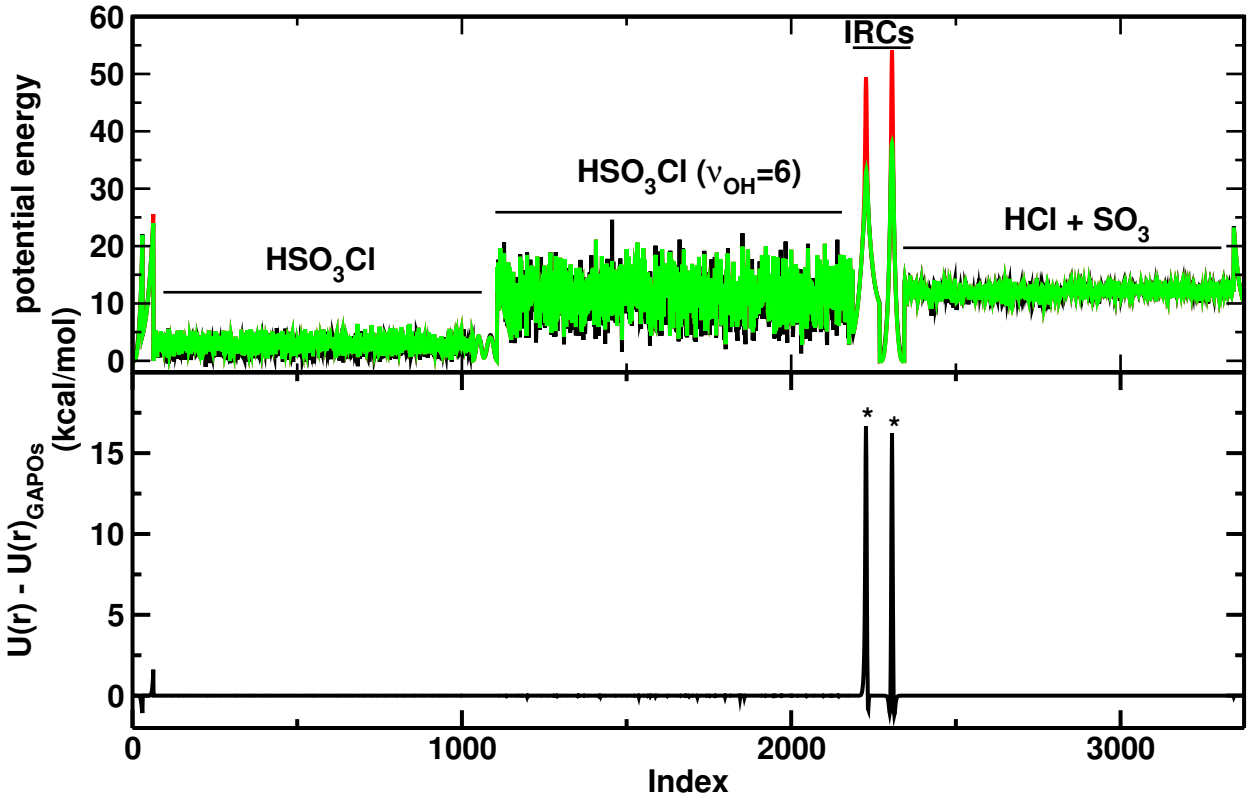


FIG. S-II: Potential energy as a function of the molecule index. Top panel) Comparison between FF with (red) and without GAPOs (green) and MP2 reference points (black). Bottom panel) Potential energy difference between FF with and without GAPOs.

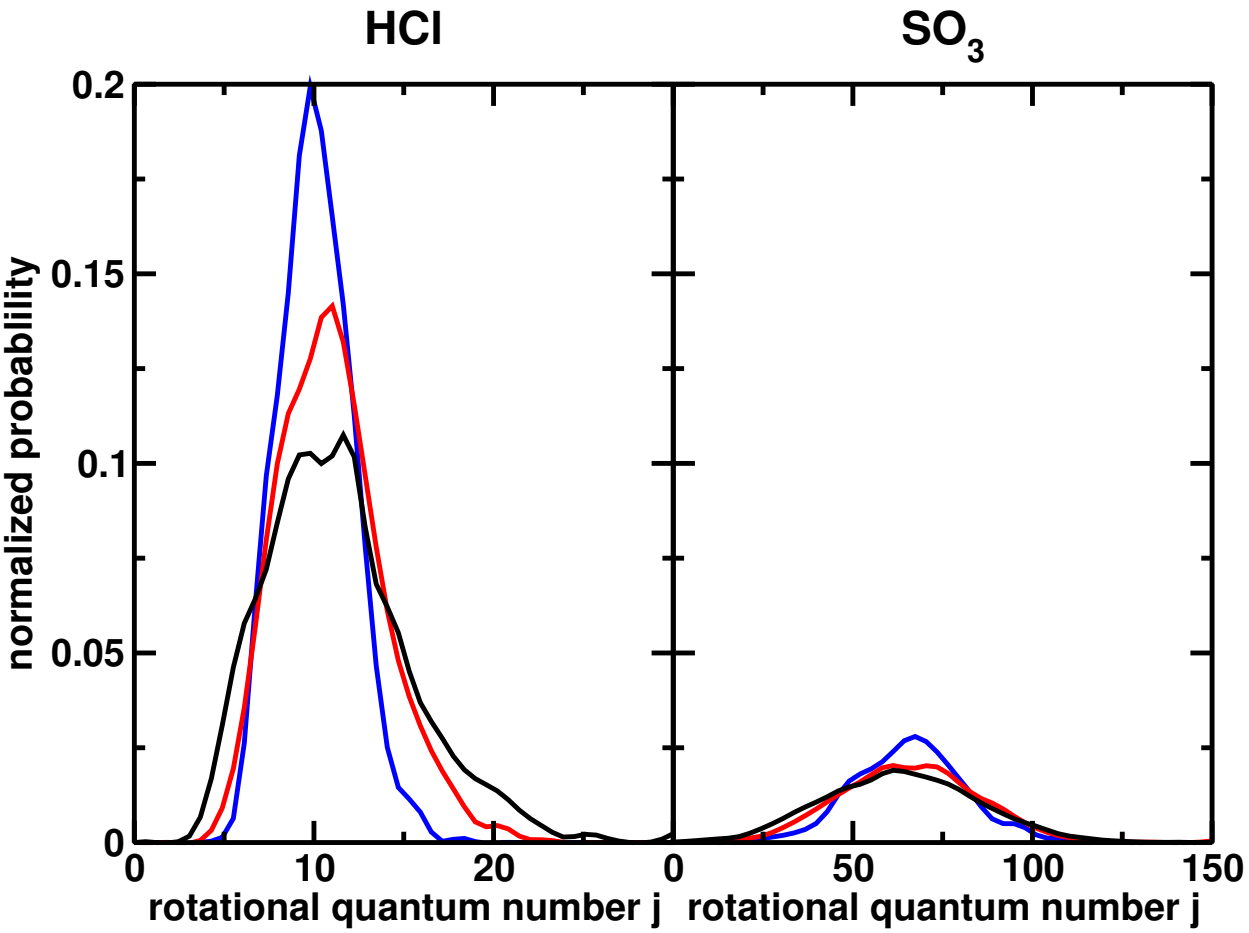


FIG. S-III: Rotational quantum number j distribution for HCl and SO_3 . Blue, red and black represents $\nu_{OH} = 4$, 5 and 6 respectively.

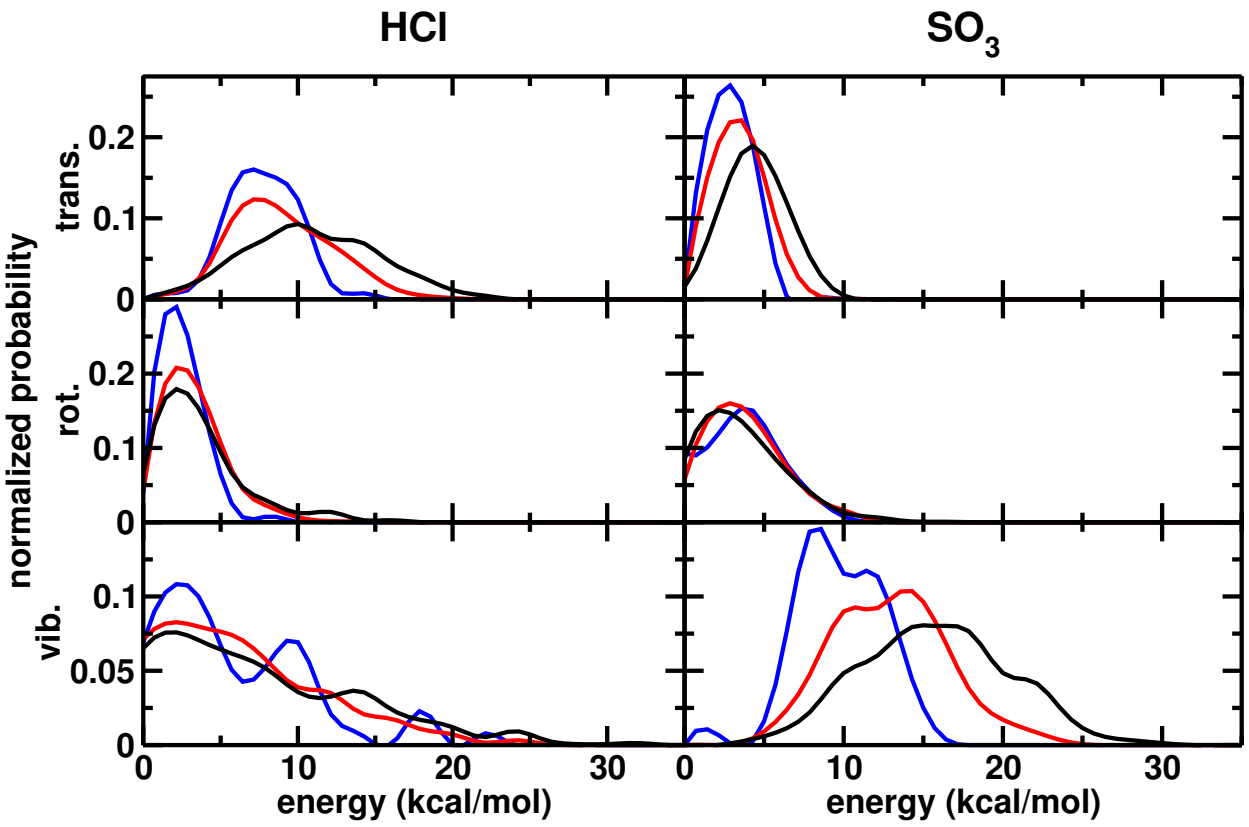


FIG. S-IV: Final state analysis for HCl and SO_3 with initial conditions from the Monte Carlo sampling. Blue, red, and black curves correspond to excitation of $\nu_{OH} = 4$, 5 and 6, respectively. See Figure 5 in the manuscript for the final state analysis from simulations with *NVE* initial conditions.

[1] J. Yosa, T. Nagy and M. Meuwly, Phys. Chem. Chem. Phys., 2014, **16**, 18533 – 18544.

[2] G. Mie, Ann. Phy., 1903, **316**, 657–697.

[3] T. Nagy, J. Yosa and M. Meuwly, J. Chem. Theory. Comput., 2014, **10**, 1366 – 1375.

[4] C. Breneman and K. Wiberg, J. Chem. Comp., 1990, **11**, 361373.

[5] V. Zoete, M. A. Cuendet, A. Grosdidier and O. Michielin, J. Comput. Chem., 2011, **32**, 2359–2368.

[6] J. Nelder and R. Mead, Chem. Phys., 1965, **7**, 308 – 313.

[7] Y. Miller and R. B. Gerber, J. Am. Chem. Soc., 2006, **128**, 9594 – 9595.

[8] P. Gupta, J. R. Lane and H. G. Kjaergaard, Phys. Chem. Chem. Phys., 2010, **12**, 8277–8284.

[9] Y. Zhao, B. J. Lynch and D. G. Truhlar, J. Phys. Chem. A, 2004, **108**, 4786–4791.

[10] A. P. Scott and L. Radom, J. Phys. Chem., 1996, **100**, 16502–16513.

3.2. Minimum Dynamic Path

The parameterized FF for HSO_3Cl was further used for the validation of the concept of a Minimum Dynamic Path (MDP).[132] MDP is defined as the path in configurational space which is followed by a trajectory that starts at the TS (with an initial momenta) with constant total energy. This concept reveals that the MEP is short of dynamical information. Furthermore, a decomposition of the total energy into its normal modes was introduced which revealed the importance of an "umbrella motion" (mode 13, see Figure 3.1) of the SO_3 moiety for the dissociation of HCl, in addition to the OH-stretch vibration (mode 18, see Figure 3.1).

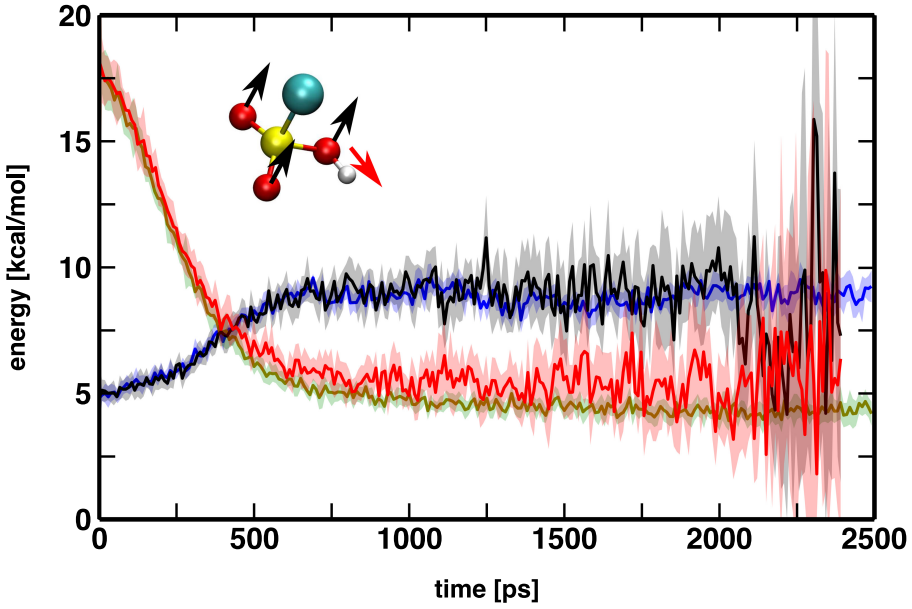


Figure 3.1.: Average normal mode energy for modes 13 (black – reactive; blue – non-reactive) and 18 (red – reactive; green – non-reactive) for 205 dissociative and 364 non-reactive HSO_3Cl trajectories after OH-stretch excitation. Shaded areas indicate one standard deviation. The inset shows the structure of HSO_3Cl and schematically with arrows the normal mode 13 (black) and 18 (red).

For this work the structure of HSO_3Cl was minimized and heated to 300 K using the velocity Verlet algorithm[128] ($\Delta t = 0.1$) during 50 ps and equilibrated for the same amount of time. Afterwards 50 ps of free dynamics simulations were performed. The only significant change to the previous work[27] was

the change of the integrator to velocity Verlet, which was necessary for the normal mode decomposition, which requires information of the coordinate and the velocities at the same time. From a total of 764 simulations with excitation of $\nu_{OH} = 5$ simulated over 2.5 ns, 205 directly dissociating trajectories (i.e. without prior H-transfer) and 364 non-reactive trajectories were analysed (the remaining trajectories showed at least one H-transfer event). Figure 3.1 shows the average decomposition of the total energy into the normal mode contributions for modes 13 and 18 for reactive and non-reactive trajectories.

It can be seen that the energy flow from mode 18 to mode 13 does not differ significantly for the two classes of trajectories. Mode 13 is excited, only prior to the reaction, as one would expect from chemical intuition. This can be traced back to the high dimensionality of the system and the resulting chaotic dynamics.

3.3. HSO_3F

As mentioned above HSO_3F was also proposed as a proxy for the overtone dissociation of sulphuric acid. Equally to HSO_3Cl, HSO_3F shows comparable barrier heights and OH-stretch overtone energies to H_2SO_4. Comparing both HSO_3X (X = F, Cl) molecules reveals dynamic bottlenecks preventing the dissociation of HSO_3F into HF and SO_3. Detailed analysis implies that coupling of the OH-local mode to bending and torsional degrees of freedom is of importance for the elimination reaction. The lack of dissociation was confirmed by *ab initio* MD simulations. The reaction probability for the H-transfer reaction in HSO_3F is high. It ranges from 75% to 98% for excitation of $\nu_{OH} = 4$ to 7. Since MS-ARMD allows for recrossing a detailed analysis of the H-transfer events is possible. The expectation values for the first H-transfer reaction indicate that the energy flow into the relevant reaction coordinates is the rate limiting step. Residence times for all subsequent events follow an exponentially decaying distribution.

HSO_3F shows strong non-RRKM (Rice-Ramsperger-Kassel-Markus) behavior. The RRKM rates for HSO_3X (X = F, Cl, Br), calculated with the Multiwell program suite[133] are shown in Figure 3.2. the analysis shows a clear trend, reflecting the size and electronegativity of X.

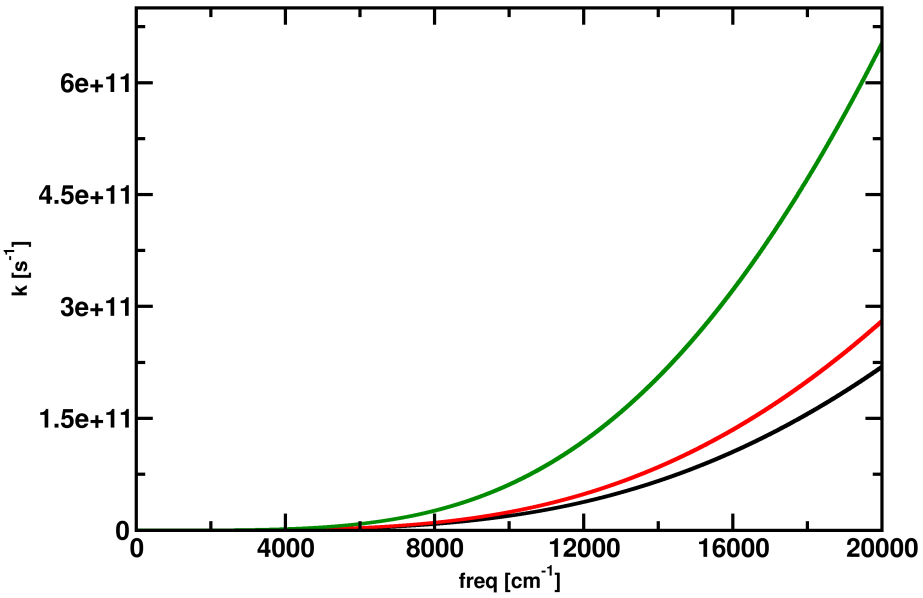


Figure 3.2.: Plot of RRKM rates of different HSO_3X (X = F: black, Cl: red, Br: green).

THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

OH-Stretching Overtone Induced Dynamics in HSO_3F from Reactive Molecular Dynamics Simulations

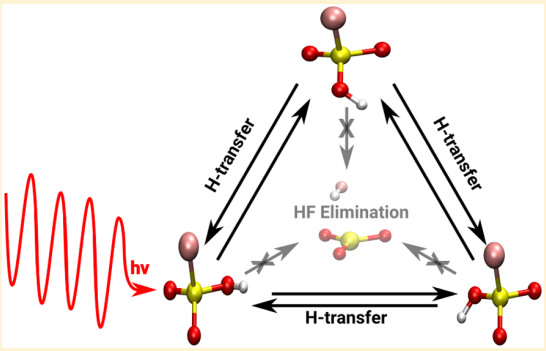
Published as part of The Journal of Physical Chemistry virtual special issue "Veronica Vaida Festschrift".

Sebastian Brickel and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

Supporting Information

ABSTRACT: The OH-stretch induced dynamics in fluorosulfonic acid (HSO_3F) is characterized from a statistically significant number of trajectories using multisurface adiabatic reactive molecular dynamics (MS-ARMD) simulations. The global reactive potential energy surface, which describes H-transfer and HF-elimination, is parametrized at the MP2/6-311G++(2p,2d) level of theory with an accuracy of better than 1 kcal/mol. Excitation along the OH-local mode leads to H-transfer dynamics but elimination of HF is only observed for excitations with $\nu \geq 6$ for 1 out of 5000 trajectories. This finding differs fundamentally from the situation for vibrationally induced photodissociation of H_2SO_4 and HSO_3Cl, for which, even with excitations of 4 quanta along the OH-stretch mode, elimination of H_2O and HCl, respectively, is readily observed on the subnanosecond time scale. RRKM rates for HX-elimination in HSO_3X (X = F, Cl) only differ by a factor of 5. The findings from the reactive molecular dynamics simulations together with the RRKM results thus indicate that the origin for a closed HF-production channel is dynamical. This is also consistent with experimental findings for hydrofluoroethanes in shock tubes, which found pronounced non-RRKM behavior.



INTRODUCTION

The study of decomposition products from energized molecules in the atmosphere is of great practical relevance. In the troposphere, interaction of stable and reactive species with sunlight leads to a myriad of decomposition products which are involved in a rich chemistry. Of particular importance are reactions and their decomposition products which lead to long-living, harmful (re)agents and those which affect the ozone layer. A topical example is the formation of sulfur-containing compounds from decomposition of H_2SO_4, which are involved in acid rain and aerosol formation in the tropo- and stratosphere and are significant for the chemistry of the atmosphere.[1] As H_2SO_4 cannot decay along a thermal and electronically driven pathway in the atmosphere, alternative reaction mechanisms needed to be considered. It was suggested[2] and later confirmed[3−6] that a vibrationally induced mechanism can drive the $H_2SO_4 \rightarrow SO_3 + H_2O$ reaction. In additional, vibrationally induced reactivity has also been investigated in mode-selective chemistry.[7,8]

Vibrational overtone induced reactions also differ in their reaction kinetics from thermally driven reactions. Population of higher vibrational states leads to a nonequilibrium preparation of the system from which it usually decays in a nonstatistical, non-RRKM (Rice−Ramsperger−Kassel−Marcus) fashion. This differs from electronic excitation that typically can be analyzed within an RRKM framework. One topical example for vibrationally induced reactivity is the photodissociation of pyruvic acid.[9,10] For this system it was demonstrated that if the decomposition reaction is initiated through pumping of the OH-stretch vibration, the kinetics does not follow an RRKM scheme, unlike the thermally induced reaction. This is important as the chemistry following a thermally or vibrationally driven process differs. The thermal process generates methylhydroxycarbene, which subsequently can isomerize to acetaldehyde. Contrary to that, vibrational overtone excitation produces relatively stable high-energy methylhydroxycarbene radicals that can drive further reactions through collisions with other collision partners, such as water.[11]

In the atmosphere, vibrationally excited molecules can undergo unimolecular reactions before vibrational quenching occurs due to collisions with surrounding material.[12−14] Especially X−H stretching (X = O, C, N) induced reactivity, which shows high vibrational frequencies compared to those of other bonds, are of significance.[12,14] A well studied example is the dissociation of HNO_x. Here the atmospheric relevance is related to releasing NO_2, which affects ozone formation.[12,15,16] Other examples of atmospherically important molecules that undergo vibrationally induced dissociation include pyruvic

Received: March 28, 2017
Revised: June 1, 2017
Published: June 15, 2017

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

acid[17] and HO_2NO_2.[18] Including vibrationally induced reactivity leads to improved atmospheric models.[12] One example is formation of atmospheric aerosols, which influences the climate, due to hydrophilic acids and alcohols in the atmosphere. Excluding vibrational overtone-induced photodissociation leads to incorrect particle sizes.[14]

The present study investigates the vibrationally induced reactivity of HSO_3F. Unlike the atmosphere of Earth, the Venusian atmosphere includes, as an important minor gas, hydrogen fluoride (HF) in its mesosphere.[19] HF is highly soluble in H_2SO_4, and small amounts of it can react with H_2SO_4, which has a concentration of about 80% in Venus' clouds,[20] to form fluorosulfonic acid (HSO_3F) and water.[21,22] It was also suggested that HSO_3F may be a proxy[13,23] for the vibrationally induced decomposition reaction of H_2SO_4, which is difficult to carry out in laboratory-based experiments. Finally, the reaction of HF with water forms hydrofluoric acid, which is known to be a common acid in polluted air[24] with negative short- and long-term effects on plant growth.[24−27]

The present computations are based on multisurface adiabatic reactive molecular dynamics (MS-ARMD[6,28,29]) simulations. MS-ARMD can treat multiple reaction pathways, on the basis of parametrized force fields (FFs) to characterize the adiabatic potential energy surface (PES). This ability is required for the description of the two different reaction pathways, accessible to HSO_3F (Figure 1). Besides the elimination pathway, leading to HF and SO_3, fluorosulfonic acid can also show intramolecular H-transfer (which can subsequently lead to HF-elimination or further H-transfer).

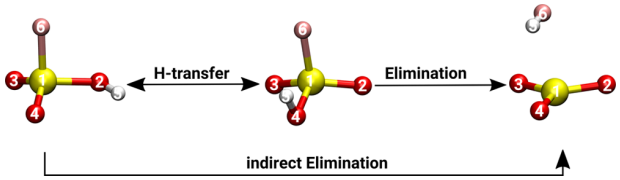


Figure 1. Graphical representation of the possible reaction paths for fluorosulfonic acid (HF-elimination, elimination; intramolecular H-transfer, H-transfer; HF-elimination after intramolecular H-transfer, indirect elimination), including atom numbering used here. Atom color code: S, yellow; O, red; F, pink; H, white.

COMPUTATIONAL DETAILS

MD Simulations. All MD simulations were carried out with CHARMM,[30] using multisurface adiabatic reactive molecular dynamics (MS-ARMD).[6] Starting from an optimized structure (10 000 steps of adopted Newton−Raphson), each system was individually heated to 300 K for 40 ps. Next, the systems were equilibrated for 40 ps at 300 K, followed by free dynamics. The time step in all simulations was Δt = 0.1 fs to conserve energy and account for the rapid hydrogen vibrational motion. The equations of motion were propagated using the leapfrog Verlet algorithm.

For the vibrational excitation, the instantaneous velocity vector was scaled along the OH-normal mode vector such as to deposit the appropriate amount of energy corresponding to a particular vibrational excitation into the molecule. After such a nonequilibrium preparation, the system was propagated in the (*NVE*) ensemble and the dynamics was analyzed in terms of (i) intramolecular hydrogen transfer from one oxygen to another one (H-transfer) or (ii) elimination of HF.

Parametrization of the Force Field. The parametrization of the force fields for each state $V_i(x)$ (i.e., connectivities), necessary for describing the reaction of HSO_3F, was accomplished by fitting the FF parameters to reference *ab initio* energies, calculated at the MP2/6-311++G(2d,2p) level of theory using Gaussian09.[31] To obtain a meaningful description of the four different states, three permutation invariant configurations of HSO_3F and one state describing the elimination products (HF and SO_3), the standard CHARMM harmonic bond potentials were replaced by Morse potentials. Furthermore in the elimination products the standard CHARMM Lennard-Jones potential between H/F and the SO_3 atoms where replaced by generalized Lennard-Jones potentials (see ref 6). The remaining FF terms were calculated as in conventional CHARMM, i.e., harmonic angle potentials in combination with Urey−Bradley terms, dihedral terms (for HSO_3F), additional improper dihedral potentials for SO_3, and point charges. After the individual parametrization of each state, the obtained FFs were combined by using energy dependent, normalized weights ($w_i(x)$) to the states $V_i(x)$. To realistically describe the shape and energetics in the crossing regions between two states, smoothing functions (Gaussian times polynomial functions, GAPOs) are used.

$$\Delta V^{ij}_{\mathrm{GAPO},k}(x) = \exp\left(-\frac{(\Delta V_{ij}(x) - V^0_{ij,k})^2}{2\sigma^2_{ij,k}}\right)\sum_{l=0}^{m_{ij,k}} a_{ij,\mathrm{kl}}(\Delta V_{ij}(x) - V^0_{ij,k})^l \quad (1)$$

They suitably modify the shape of the reactive potentials on the basis of the difference of the potential energy of two states ($\Delta V_{ij} = V_i(x) + V_j(x)$) to reproduce the energies from *ab initio* calculations (i.e., IRC calculations) in the barrier region. The center and width of the Gaussian function are given by $V^0_{ij,k}$ and $\sigma^2_{ij,k}$, respectively. The order of the polynomial is given by $m_{ij,k}$; $a_{ij,kl}$ is the coefficient of the polynomial. The global PES is then given by the sum of the weighted sum of PESs and the weighted sum of the GAPOs.

$$V_{\mathrm{MS\text{-}ARMD}}(x) = \sum_{i=1}^{n} w_i(x)\, V_i(x) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} [w_i(x) + w_j(x)] \sum_{k=1}^{n_{ij}} \Delta V^{ij}_{\mathrm{GAPO},k}(x) \quad (2)$$

Parametrization of the individual states $V_i(x)$ started with initial parameters obtained from SwissParam.[32] Representative structures of the reactant and the product were sampled from 500 ps MD simulations at 300 K. For these structures, reference energies were determined at the MP2/6-311++G(2d,2p) level of theory. The fitting of the individual FFs for each state was performed using a downhill simplex algorithm.[33] For the parametrization of the GAPOs the IRCs of the two possible reaction mechanisms (hydrogen transfer and HF-elimination) were also calculated at the MP2/6-311++G(2d,2p) level of theory. For fitting the parameters of the GAPOs a genetic algorithm was used, as their functional form leads to a highly nonlinear fit.

RESULTS AND DISCUSSION

Quality of the Force Field. The parametrization of the reactant and product states was performed in an iterative fashion. Starting with a set of 100 structures for HSO_3F and

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

SO_3 + HF a first fit for the two states was carried out. This was followed by further MD simulations using this improved set of parameters from which another 500 structures was extracted and included in the fit. Following several similar rounds of refinements, the fit of the individual states included 5200 reference points in total. The root mean squared deviation for the final set between the target and the fitted energies was 0.93 and 0.58 kcal/mol, respectively, for the HSO_3F and SO_3 + HF state (blue and green points in Figure 2). The energies of the

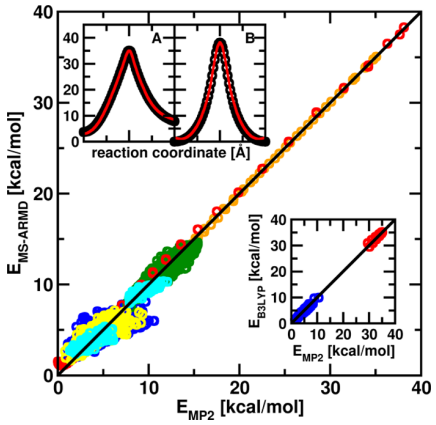


Figure 2. Energy correlation of 4865 MP2/6-311++G(2d,2p) reference structures and the fitted FF. Color code: blue = HSO_3F, green = SO_3 + HF, orange = IRC elimination, red = IRC H-transfer, yellow = 500 structures of HSO_3F from 100 ns simulation, cyan = additional HSO_3F and SO_3 + HF. Top left inset shows the *ab initio* IRCs (black circles, calculated with MP2/6-311++G(2d,2p)) and fitted FF (red curves, including fitted GAPOs). (A) HF-elimination. (B) Intramolecular H-transfer. Bottom right inset shows the energy correlation for 200 structures of HSO_3F (blue) and 11 points around the MP2 elimination IRC calculated with B3LYP/6-311+G(2d,2p) and MP2/6-311++G(2d,2p).

IRC structure were calculated with these states $V_i(x)$ and provided to the GAPO fit procedure. For the parametrization of the elimination reaction, a RMSD of 0.30 kcal/mol was achieved (Figure 2A) using three Gaussians, one with polynomial order up to four and two with polynomials up to order five. For H-transfer three pure Gaussian GAPOs yield a RMSD of 0.63 kcal/mol (Figure 2B).

For validation, the minimized structures of the reactant and product states were determined from the fitted MS-ARMD force field and compared with those from the reference electronic structure calculations. The aligned structures have an RMSD of 0.02 and 0.14 Å, respectively, for HSO_3F and SO_3 + HF. As a second comparison, the normal modes were determined. Their average RMSDs for HSO_3F is 46.7 cm^{-1} and for SO_3 + HF 70.0 cm^{-1}, which is a reasonable agreement for our purposes. As a final validation, the energies for 500 structures from a 100 ns dynamics were determined from the fitted force field and from MP2/6-311++G(2d,2p) calculations. The RMSD between reference and target values is 0.9 kcal/mol, which underlines the quality of the force field.

Intramolecular H-Transfer. First, intramolecular hydrogen transfer is considered. For this, 5000 independent trajectories, each 2.5 ns in length were run for OH-excitations with ν_{OH} = 4, 5, 6, or 7 quanta. The energies of the anharmonic OH-vibrations for ν_{OH} = 4−8 are summarized in Table 1.

Table 1. Calculated OH Bond Stretching Frequencies of HSO_3F (kcal/mol)[a]

ν_{OH}	MS-ARMD	CC-VSCF on MP2/TZP[13]
4	41.41	39.06
5	50.43	49.53
6	59.14	61.47
7	67.52	75.20
8	75.59	N/A

[a]MS-ARMD values are compared to CC-VSCF frequencies calculated at the MP2/TZP level of theory.[13]

The reaction probabilities ranged from 75% for ν_{OH} = 4 to 98% for all higher excitations. For ν_{OH} = 4, 33% of the trajectories showed only one single H-transfer within 2.5 ns after which intramolecular energy redistribution precluded further reactions. Excitations with 5 and more quanta always lead to multiple hydrogen transfers during the 2.5 ns.

MS-ARMD follows reactions explicitly in time and hence recrossing after transit of the transition state is possible. Such recrossing is also observed here. When residence times of the hydrogen atom on a particular oxygen are analyzed, events with lifetimes shorter than 500 fs are excluded. Residence times τ following vibrational excitation and those after the first hydrogen transfer has occurred were analyzed separately. The expectation values ($\langle\tau\rangle = \int \tau p(\tau)\, d\tau$) for the first hydrogen transfer to occur is 890.6, 449.3, 117.3, and 55.0 ps for excitations with ν_{OH} = 4 to 7. After the first transfer, the average residence times are 357.8, 89.9, 27.6, and 12.0 ps, respectively. The residence time distributions $p_1(\tau)$ for excitation with ν_{OH} = 5 before the first transfer and $p_{>1}(\tau)$ for all subsequent H-transfer events are shown in Figure 3. Here it is found that the first H-transfer occurs after ≈100 ps with a broad distribution extending up to 1.7 ns. After the first transfer, the residence times follow an exponentially decaying

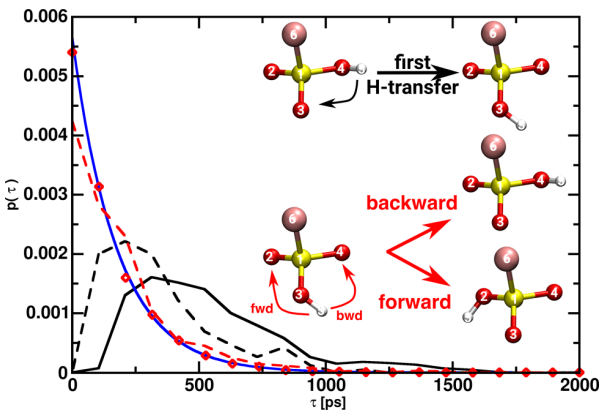


Figure 3. Normalized residence probability distribution $p_1(\tau)$ before the first H-transfer takes place for HSO_3F (black solid) and for HSO_3Cl (black dashed). Raw data for the distribution of all subsequent H-transfer events $p_{>1}(\tau)$ for HSO_3F (red diamonds) follow a single exponential decay (blue solid line) and raw data for HSO_3Cl (red dashed) are similar but the distribution is wider and is shifted to somewhat longer time scales. The reaction in the top right corner shows one of the two possible first H-transfer reactions (from O4 to O3). The subsequent H-transfer reaction, which was considered for determination of the residence time, is shown in the bottom right corner. Backward (bwd) reaction from O3 to O4 and forward (fwd) reaction from O3 to O2.

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

distribution (red trace in Figure 3). This is likely to be related to the fact that once H-transfer occurred (i.e., the necessary energy is in the relevant reaction coordinate) the probability for back-transfer is large (≈60% of the events). For the first H-transfer to occur the efficiency of energy flow into the relevant reaction coordinate is the rate limiting step, which depends directly on the initial conditions.

Figure 3 shows the $p(\tau)$ and the residence time distribution for HSO_3Cl when $\nu = 5$ is excited. The first hydrogen transfer event in HSO_3Cl takes place immediately after excitation and shows a slight secondary maximum around 700 ps. The average residence times before the first transfer for excitation with ν = 4−6 are 573.5, 242.1 (black, dashed line in Figure 3), and 65.2 ps for HSO_3Cl, which are shorter by close to a factor of 2 compared to those for HSO_3F despite a similar barrier height for the H-transfer reaction. After the first transfer has occurred, residence times are slightly longer for HSO_3Cl than for HSO_3F.

HF-Elimination Reaction. Next, HF-product formation was considered. For any of the excitations with ν_{OH} = 4, 5, and 6 none of the 5000 independent trajectories showed elimination. Excitation levels of ν_{OH} = 4−6 correspond to an increase in the internal energy between 40 and 60 kcal/mol, which is considerably higher than the barrier for HF-elimination, which is 35.1 kcal/mol. As mentioned above, most trajectories exhibit H-transfer, which is a potential channel for efficient energy redistribution such that recurrence times can become very long. To validate this, 10 trajectories with ν_{OH} = 6 were extended up to 100 ns. However, none of these trajectories showed HF-elimination. For a broader assessment, 10^5 independent trajectories were run from initial structures generated from a Monte Carlo simulation where structures were saved every 20 000 MC steps. After a short heating, equilibration and dynamics (each 10 ps) MS-ARMD simulations with ν_{OH} = 6 were performed in the same manner as explained above. None of these trajectories showed elimination within 100 ps.

In an attempt to force HF-elimination, 5000 additional trajectories were run with excitation of ν_{OH} = 7 and 500 trajectories with excitation ν_{OH} = 8, respectively. Product formation was found in five and eight cases for the two excitation levels. Hence, the reaction probability is excessively small despite the large energy deposited compared with the barrier of 35.1 kcal/mol for the elimination reaction. This points toward the importance of couplings between internal degrees of freedom to move from reactant to product in the present case, a point that will be considered further below.

The finding that the reaction probability for HF-elimination is negligible prompted us to consider the possible reasons in some more detail. First, the minimum energy and transition state geometries for the HX-elimination reaction are compared. One notable difference between the two transition state structures is the OH bond length, which is 1.21 Å for HSO_3F compared to 1.27 Å for HSO_3Cl (Figure 4). This suggests that the chlorinated complex has a weakened OH-bond compared to the fluorinated species. The other coordinates either are similar (such as valence angles) or can be explained by the larger size of Cl vs F.

Next, the reaction dynamics in HSO_3X (X = F and Cl) was compared. First, it is noted that even for moderate excitation levels (ν_{OH} = 4) along the OH-stretch coordinate HCl-elimination in HSO_3Cl occurs for 15% of the trajectories on a 2.5 ns time scale.[34] This compares with no elimination at all for HSO_3F for the same level of excitation and over the same time

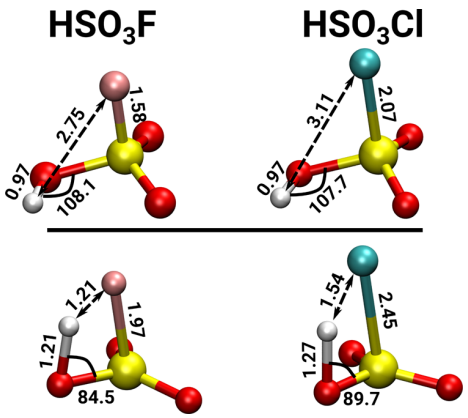


Figure 4. *Ab initio* (MP2/6-311++G(2d,2p)) geometry optimized structures (top) and barrier geometries leading to the elimination reaction (bottom) for HSO_3X (X = F, left, Cl, right). The barrier heights for HX-elimination and H-transfer are 35.1 kcal/mol vs 33.1 kcal/mol and 38.1 kcal/mol vs 38.1 kcal/mol for HSO_3F and HSO_3Cl, respectively.

interval after vibrational excitation. The computed MP2/6-311++G(2d,2p) barriers for HX-elimination and H-transfer for the two systems are 35.1 kcal/mol vs 33.1 kcal/mol and 38.1 kcal/mol vs 38.1 kcal/mol for HSO_3F and HSO_3Cl, respectively.[34]

Given the similar barrier heights for the two systems, it is meaningful to explicitly compare the molecular dynamics of trajectories that do and do not show HX-elimination, respectively. To this end, 2-dimensional probability distributions from trajectories with vibrationally excited (ν_{OH} = 5) HSO_3F and HSO_3Cl were considered; see Figures 5 and 6 for several combinations of internal degrees of freedom. This particular level of excitation was chosen because for HSO_3Cl a large fraction of trajectories (more than 90%) show HCl-elimination and, at the same time, a statistically sufficient number of nondissociative trajectories exists from which probability distributions can be constructed.

Comparison of the normalized probability distributions $P(\phi,r_{SX})$, $P(\phi,r_{OH})$, $P(\phi,\angle XSO)$, and $P(\phi,\angle SOH)$ (where ϕ is the X−S−O−H dihedral angle) from simulations for HSO_3F and for HSO_3Cl points toward the importance of several internal degrees of freedom that are involved and govern HX-elimination (where X is Cl or F, respectively). The first, second, and third rows in Figure 5 show the probability distributions for HSO_3F and HSO_3Cl without elimination. The fourth row in Figure 5 shows P for HSO_3Cl with HCl-elimination. These distributions were generated from data 1 ps before reaching the transition state and including a few time steps toward the product. The panels showing P for HSO_3Cl also report the distributions of TS geometries for dissociative trajectories (without prior H-transfer).

The probability distributions of these five degrees of freedom (Figure 5) show major differences for dissociative and nondissociative trajectories (compare rows 3 and 4 for HSO_3Cl), respectively. This serves as a suitable proxy to discuss the situation in HSO_3F. The red point clouds in Figure 5 are the ensemble of crossing points for trajectories with HCl-elimination in HSO_3Cl. As is evident, noneliminating trajectories do not sample these regions in configuration space whereas those leading to the product evidently do sample the crossing geometries for HSO_3Cl. Interestingly, the

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

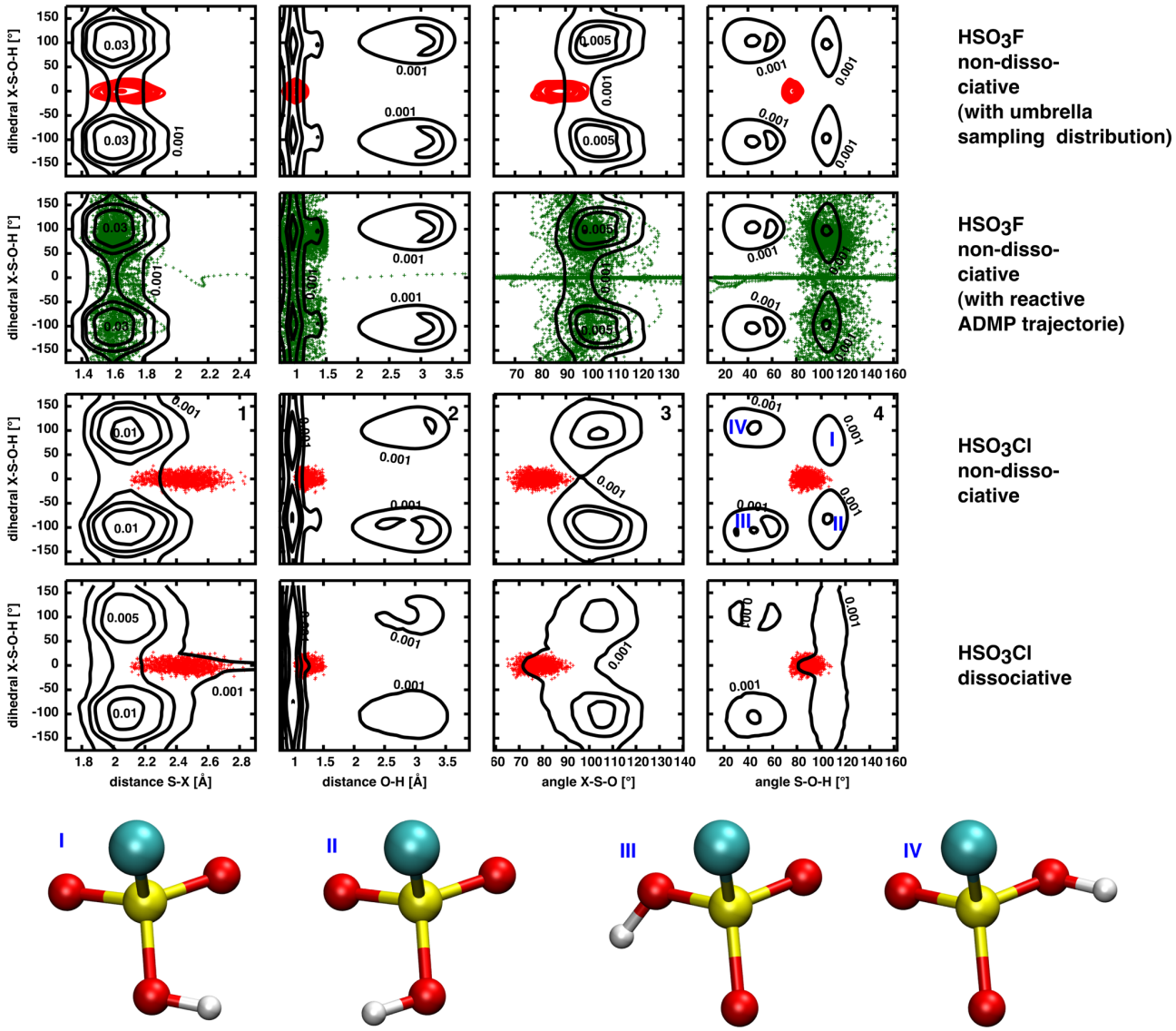


Figure 5. Internal coordinate probability distributions (normalized to unity for each panel and the lowest isocontour labeled in each case). From left to right: $P(\phi,r_{SX})$, $P(\phi,r_{OH})$, $P(\phi,\angle XSO)$, and $P(\phi,\angle SOH)$, with ϕ = X−S−O−H dihedral angle where X = F, Cl. From top to bottom: coordinate probability distributions (black) from nondissociative trajectories for HSO_3F together with probability distributions for structures extracted from umbrella sampling with ρ = 1.685 (red); same as top row but with structures sampled in a dissociative *ab initio* MD trajectory of HSO_3F (green); nondissociative HSO_3Cl trajectories and dissociative HSO_3Cl trajectories 1 ps before reaching a S−Cl distance of 3 Å. The red crosses in rows three and four mark the TSs of dissociative HSO_3Cl trajectories. Bottom: Structures I and II show the two synperiplanar positions of H with respect to $O_{initial}$. III and IV show the two different configurations after intramolecular H-transfer reaction.

probability distribution functions for HSO_3F resemble very closely those for noneliminating HSO_3Cl. These distributions suggest that the relevant degrees of freedom are those sampled in the dissociative trajectories but not in the nondissociative trajectories. These are the S−O−H angle and the X−S−O−H dihedral angle. Additionally, a minor contribution arises from the F−S−O angle.

The difference in configuration space sampled by nondissociative and dissociative HSO_3Cl trajectories can also be seen from the TS distribution of dissociative HSO_3Cl trajectories (red crosses in Figure 5). The TS has a coplanar structure (dihedral angle Cl−S−O−H ≈ 0°; see bottom structures in Figure 4). Furthermore, it is noticeable that for dissociation the S−O−H angle needs to be smaller than its equilibrium value of 110° (HSO_3Cl panel 4, Figure 5). Likewise, the Cl−S−O angle takes smaller values in the TS than in the reactant conformation (HSO_3Cl panel 3, Figure 5).

The bottom of Figure 5 shows some notable HSO_3Cl structures that are sampled for nondissociating trajectories. The two different synperiplanar conformations of the reactant (I and II) as well as the two different H-transfer product configurations. Panels 1 and 3 in all four rows in Figure 5 show two maxima that simultaneously sample configuration III (in both conformations) together with I and configuration IV (in both conformations) together with II. In panels 2 and 4 in all rows the three configurations can be identified. In panels 2 the distributions at short O−H distances correspond to I and II, those at longer O−H distances (>1.5 Å) correspond to III and

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

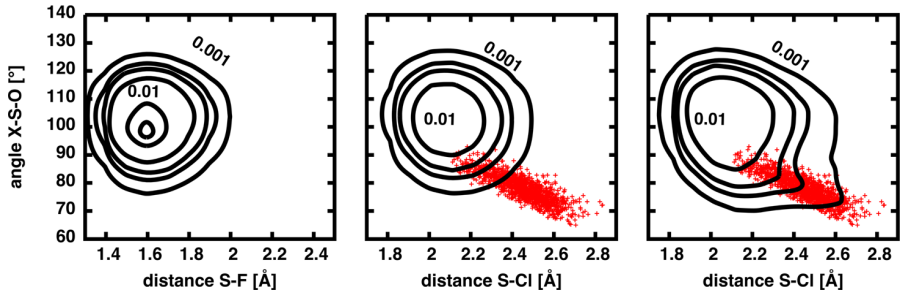


Figure 6. Probability distributions $P(\angle XSO, r_{SX})$ (normalized to unity for each panel and the lowest isocontour labeled in each case), where X = F, Cl. From left to right: nondissociative HSO_3F trajectories, nondissociative HSO_3Cl trajectories, and dissociative HSO_3Cl trajectories 1 ps before reaching a S−Cl distance of 3 Å. Red crosses in the middle and right panels mark the TSs of dissociative HSO_3Cl trajectories.

IV separated by their different dihedral angle distribution. Furthermore, in panel 2 the two synperiplanar conformations I and II can be identified by the X−S−O−H dihedral angle (X = F, Cl). This is similar for panel 4, where small S−O−H angles correspond to structures III and IV, whereas the two maxima for bigger S−O−H angles correspond to I and II.

The probability distributions of $P(\angle XSO, r_{SX})$ with X being F and Cl, respectively, are shown in Figure 6. The nondissociative trajectories of HSO_3F and HSO_3Cl are circular, whereas dissociative trajectories of HSO_3Cl sample longer S−Cl distances in combination with shorter Cl−S−O angles, therefore tending toward the TSs. Structurally, the shorter angle decreases the distance between Cl and H, while the longer S−Cl distance favors the S−Cl bond breaking. These conformations appear to be inaccessible in HSO_3F.

From these probability distributions (Figures 5 and 6) it can be seen, by comparing the phase space sampled, that coupling between the OH-stretch and SOH-bend, XOS-bend, SX-stretch, and the torsion XSOH (X= F, Cl) is crucial for OH-stretching overtone induced photodissociation.

To better characterize the region around the transition state, umbrella sampling simulations were carried out for the HF-elimination reaction.[35] These simulations used the ratio $\rho = r_{SF}/r_{FH}$ as the reaction coordinate. The reaction coordinate was scanned from ρ = 0.5 to ρ = 3.0, and the data were combined using the weighted histogram analysis method (WHAM).[36] The barrier is 38.5 kcal/mol, compared with a forward activation energy of 35.1 kcal/mol. The transition state (TS) is located at ρ = 1.685, i.e., for ρ < 1.685 the system is in its reactant state whereas for ρ > 1.685 the product is formed. To verify its nature, an umbrella sampling simulation 12.5 ns in length was carried out and 1750 representative structures were used as starting points for energy minimization which were free to evolve toward product or reactant, respectively. The fraction of structures relaxing toward reactant and product was 57% and 43%, respectively, which suggests that the ensemble characterizes the TS in a meaningful way. However, further refinement to obtain the point with a committor probability of 0.5 would, in principle, be possible but was not pursued here. The averaged structure over these 1750 samples is reported in the inset in Figure 7. The normalized probability distributions of $P(\phi, r_{SF})$, $P(\phi, r_{OH})$, $P(\phi, \angle FSO)$, and $P(\phi, \angle SOH)$ for all 1750 samples are shown in Figure 5 (red distribution in the first row).

Comparison of the distributions $P(\angle FSO, \angle FSOH)$ and $P(\angle SOH, \angle FSOH)$ from HSO_3F (red distribution in the first row) and dissociative HSO_3Cl (fourth row) in Figure 5 shows that the sampling of the angle SOH is around 10° smaller in HSO_3F compared to that for HSO_3Cl. The harmonic frequencies of the SOH bending and the XSOH torsion in the two systems are 1149 cm^{-1} vs 1141 cm^{-1} and 219 cm^{-1} vs 258 cm^{-1} for HSO_3F and HSO_3Cl, respectively.[34] In addition, the Fourier transform of the power spectrum for these excitations in HSO_3F and HSO_3Cl were calculated from 100 ps of equilibrium MD simulations. Anharmonic frequencies for the SOH bending of HSO_3F and HSO_3Cl were found to be at 1157 and 1148 cm^{-1}, respectively. The XSOH torsion is 307 and 353 cm^{-1} for HSO_3F and HSO_3Cl.

Additional insight into the reaction mechanism may be obtained from an RRKM analysis of the HX-elimination pathway for both HSO_3X (X = F and Cl) systems. For this, harmonic frequencies for the global minimum (Table S1) and the TS for HSO_3F and HSO_3Cl were determined at the MP2/6-311++(2d,2p) level of theory and RRKM rates were computed by using the MultiWell program.[37] The potential energies of the dissociated products relative to the reactants are 7.99 and 7.89 kcal/mol for HSO_3F and HSO_3Cl, respectively. A collision free regime was assumed and tunneling was neglected. For the energy range equivalent to an excitation with ν = 4−6 it was found that the RRKM rates for HSO_3F are only a factor of 5 smaller than the rate for HSO_3Cl. Such a small difference indicates that the differences for HX-elimination in HSO_3F and HSO_3Cl found in the reactive MD simulations are not predominantly a statistical but rather a dynamical effect.

Finally, the coupling between different internal modes was probed from simulations started at the transition state of the elimination reaction with zero temperature and evolving toward the reactant. The power spectrum along the OH-stretching coordinate indicates that the number of modes to which this degree of freedom couples is considerably larger in HSO_3Cl than in HSO_3F. This suggests that coupling to softer degrees of freedom in HSO_3F is less efficient, which renders transfer of the OH-stretch energy into modes that promote the system toward the HF-elimination product more difficult.

***Ab Initio* MD.** For a final validation of the findings concerning a small probability for HF-elimination, *ab initio* MD simulations were carried out. For this, "Atom Centered Density Matrix Propagation" (ADMP[38,39]) molecular dynamics as implemented in Gaussian09 was employed. ADMP uses an extended Lagrangian approach and Gaussian basis functions for propagation of the density matrix.

The simulations were carried out at the B3LYP/6-311+G-(2d,2p) level of theory. Coordinates and velocities (after overtone excitation with ν_{OH} = 6) were taken from the ensemble of structures utilized in the MS-ARMD simulation. The level of excitation was chosen to increase the probability to

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

observe elimination within a few picoseconds of simulation time. The temperature was set to 300 K.

To validate the use of the B3LYP/6-311+G(2d,2p) level of theory, the energies for 200 representative HSO_3F structures around the reactant and transition state were calculated and compared with energies from MP2/6-311++G(2d,2p) calculations. Furthermore, the energy of the TS from the IRC calculation from MP2, as well as five points before and five points after the TS along the IRC path were included in the comparison. The RMSD between B3LYP and MP2, both with 6-311+g(2,2p), is 0.4. The regression coefficient of the linear fit is 0.94. The energy correlation between both methods is shown in the bottom right inset in Figure 2. It is found that B3LYP/6-311+G(2d,2p) is a meaningful level of theory for ADMP simulations.

ADMP simulations were carried out for 20 ps with a time step of Δt = 0.2 fs. Only 1 out the 43 trajectories exhibited HF-elimination after 9.5 ps. The trace of this trajectory (dark green) in terms of the HF and SF distances sampled is shown in Figure 7. This can be compared with HF and SF distances

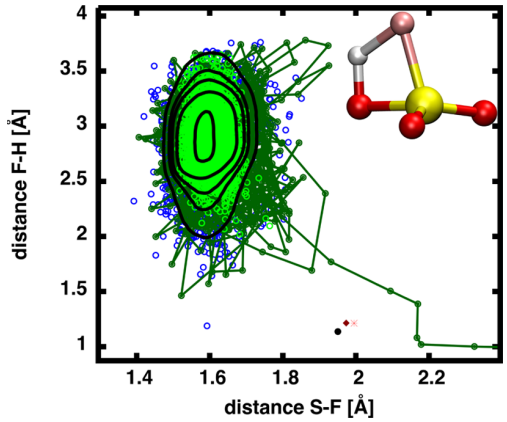


Figure 7. Normalized F−H and S−F distance distribution from 4891 MS-ARMD simulations with ν_{OH} = 6 (black lines, lowest isocontour labeled). Red dots mark the position of the transition state (red, B3LYP; dark red, MP2). Blue dots show the phase space sampled by a MS-ARMD trajectory with ν_{OH} = 6 (2.5 ns), green dots show the distribution of an *ab initio* MD trajectory that shows no reaction (within 20 ps), and dark green points combined by lines the distribution of the *ab initio* MD trajectory that with HF-elimination within 20 ps. The inset shows the average structure of the umbrella sampling TS.

sampled by an ADMP trajectory without elimination (light green) and an MS-ARMD trajectory without elimination (blue) for the same degree of excitation. As is evident, when the HF and SF distances are considered, no noticeable differences between the three trajectories are found. Figure 7 furthermore shows the normalized F−H and S−F distance distribution from 4891 MS-ARMD simulations with ν_{OH} = 6 (black lines). The three traces in Figure 7 lie well within the distribution of the ν_{OH} = 6 MS-ARMD trajectories.

This can be compared with HF and SF distances sampled by an ADMP trajectory without elimination (light green) and an MS-ARMD trajectory without elimination (blue) for the same degree of excitation. As is evident, when the HF and SF distances are considered, no noticeable differences between the three trajectories are found. The trace of the dissociative trajectory is also shown in Figure 5 (green crosses in row two).

Discussion and Conclusion. Vibrationally induced reactions are important in a number of circumstances as they differ in fundamental ways from thermally induced reactions. In particular, they can exhibit mode specificity and nonstatistical behavior. Decomposition reactions following vibrational OH-excitation have been found for species such as H_2SO_4,[5,40−44,44−47] HONO,[15,48−50] HOONO,[51] or HSO_3Cl.[13,23,42] In all these cases it is known that energies between 30 and 60 kcal/mol are required to drive the reaction, which implies that higher overtones need to be excited if a vibrationally driven process is envisaged. However, due to their small cross sections (3−6 orders of magnitude smaller than electronic transitions at the same energy[12]), product yields in laboratory experiments from vibrationally induced reactions are small, almost independent of the quantum yield. In the atmosphere, the photodissociation rate depends not only on the (wavelength dependent) cross section and the quantum yield but also on the actinic flux which is, however, almost constant (at 10^{14} photons cm^{-2} s^{-1}) over the relevant energy range for vibrational excitations of 4−6 OH-quanta[12] but may vary with altitude.[52]

The HF-yield from OH-stretch excitation of HSO_3F is found to be negligible (1 out of 5000 trajectories for $\nu_{OH} \geq 6$). This sets it apart from similar (and related) systems such as H_2SO_4[47] or HSO_3Cl[34] using the same approach. For both, H_2SO_4 and HSO_3Cl, it was found that after 500 ps more than 50% of the trajectories lead to H_2O- or HCl-elimination, respectively, for ν_{OH} = 6. Together with the observation that the excitation cross section for vibrational overtones are intrinsically small and decrease by approximately 1 order of magnitude with each quantum of excitation,[53] HF-production from OH-stretch induced photodissociation of HSO_3F is suppressed compared to the cases for H_2SO_4 and HSO_3Cl. Hence, HSO_3F is not a meaningful proxy in laboratory studies to study vibrationally induced photodissociation of compounds such as H_2SO_4.

The present findings of pronounced non-RRKM behavior are also consistent with earlier studies that considered HF-elimination from hydrofluoroethanes in shock tube experiments at elevated temperatures (1600−2400 K).[54,55] In particular, the decomposition of $CF_3CH_3 \rightarrow CH_2CF_2 + HF$ has been considered to be "the most unambiguous example of an "intrinsic" non-RRKM unimolecular process".[54]

The present analysis also suggests that coupling between radial and softer (torsional and angular) degrees of freedom is essential for promoting an elimination reaction in such systems. This is consistent with previous *ab initio* MD simulations (PM3) that found the excitation along the OH local mode yields HF-elimination in 0 to 1 (out of 32) trajectories for ν_{OH} = 5−7.[13] The yield increases to 7 up to 25 (out of 48) if 1 quantum of SOH bending is excited together with excitation of the OH-stretch. However, exciting 1 quantum of the O−S−O−H torsion together with the OH-stretch also increases the number of trajectories with HF-elimination (9 to 21 out of 48). Finally, these results are also consistent with the observation that in elevated temperature (≈400 K), near-IR gas phase measurements for HSO_3F are more stable than those for HSO_3Cl.[56]

In conclusion, the present work shows that HF-production from OH-stretching overtone induced photodissociation of HSO_3F is insignificant. Following vibrational excitation with 4 and more quanta along the OH local mode, intramolecular H-transfer is the dominant reaction channel. A comprehensive characterization of vibrationally induced HF-elimination in

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

HSO_3F can be carried out due to exhaustive sampling, which is possible because MS-ARMD simulations allow us to run a statistically significant number of trajectories for sufficiently long simulation times. The analysis points toward the importance of the coupling between radial and angular/dihedral degrees of freedom to promote reactivity in such systems.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.7b02950.

MP2 frequencies (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*M. Meuwly. E-mail: m.meuwly@unibas.ch. Phone: +41 (0)61 207 38 21.

ORCID

Sebastian Brickel: 0000-0003-1972-0393

Markus Meuwly: 0000-0001-7930-8806

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation through grants 200021-117810, the NCCR MUST, and the University of Basel.

■ REFERENCES

(1) Finlayson-Pitts, B. J. Halogens in the Troposphere. *Anal. Chem.* **2010**, *82*, 770−776.

(2) Vaida, V.; Kjaergaard, H. G.; Hintze, P. E.; Donaldson, D. J. Photolysis of Sulfuric Acid Vapor by Visible Solar Radiation. *Science* **2003**, *299*, 1566−1568.

(3) Miller, Y.; Gerber, R. B. Dynamics of Vibrational Overtone Excitations of H_2SO_4, H_2SO_4-H_2O: Hydrogen-Hopping and Photodissociation Processes. *J. Am. Chem. Soc.* **2006**, *128*, 9594−9595.

(4) Miller, Y.; Gerber, R. B.; Vaida, V. Photodissociation Yield for Vibrationally Excited States of Sulfuric Acid Under Atmospheric Conditions. *Geophys. Res. Lett.* **2007**, *34*, 1−5.

(5) Yosa Reyes, J.; Meuwly, M. Vibrationally Induced Dissociation of Sulfuric Acid (H_2SO_4). *J. Phys. Chem. A* **2011**, *115*, 14350−14360.

(6) Nagy, T.; Yosa Reyes, J.; Meuwly, M. Multisurface Adiabatic Reactive Molecular Dynamics. *J. Chem. Theory Comput.* **2014**, *10*, 1366−1375.

(7) Crim, F. F. Chemical Dynamics of Vibrationally Excited Molecules: Controlling Reactions in Gases and on Surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 12654−12661.

(8) Hundt, P. M.; Jiang, B.; van Reijzen, M. E.; Guo, H.; Beck, R. D. Vibrationally Promoted Dissociation of Water on Ni(111). *Science* **2014**, *344*, 504−507.

(9) Plath, K. L.; Takahashi, K.; Skodje, R. T.; Vaida, V. Fundamental and Overtone Vibrational Spectra of Gas-Phase Pyruvic Acid. *J. Phys. Chem. A* **2009**, *113*, 7294−7303.

(10) Vaida, V. Spectroscopy of Photoreactive Systems: Implications for Atmospheric Chemistry. *J. Phys. Chem. A* **2009**, *113*, 5−18.

(11) Takahashi, K.; Plath, K. L.; Skodje, R. T.; Vaida, V. Dynamics of Vibrational Overtone Excited Pyruvic Acid in the Gas Phase: Line Broadening Through Hydrogen-Atom Chattering. *J. Phys. Chem. A* **2008**, *112*, 7321−7331.

(12) Vaida, V.; Donaldson, D. J. Red-Light Initiated Atmospheric Reactions of Vibrationally Excited Molecules. *Phys. Chem. Chem. Phys.* **2014**, *16*, 827−836.

(13) Gupta, P.; Lane, J. R.; Kjaergaard, H. G. Molecular Dynamic Simulations of OH-Stretching Overtone Induced Photodissociation of Fluorosulfonic and Chlorosulfonic Acid. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8277−8284.

(14) Vaida, V.; Feierabend, K. J.; Rontu, N.; Takahashi, K. Sunlight-Initiated Photochemistry: Excited Vibrational States of Atmospheric Chromophores. *Int. J. Photoenergy* **2008**, *2008*, 1−13.

(15) Gerber, R. B.; Sebek, J. Dynamics Simulations of Atmospherically Relevant Molecular Reactions. *Int. Rev. Phys. Chem.* **2009**, *28*, 207−222.

(16) Salawitch, R. J.; Wennberg, P. O.; Toon, G. C.; Sen, B.; Blavier, J. F. Near IR Photolysis of HO_2NO_2: Implications for HO_x. *Geophys. Res. Lett.* **2002**, *29*, 9-1−9-4.

(17) Mellouki, A.; Mu, Y. On the Atmospheric Degradation of Pyruvic Acid in the Gas Phase. *J. Photochem. Photobiol., A* **2003**, *157*, 295−300.

(18) Stark, H.; Brown, S. S.; Burkholder, J. B.; Aldener, M.; Riffault, V.; Gierczak, T.; Ravishankara, A. R. Overtone Dissociation of Peroxynitric Acid (HO_2NO_2): Absorption Cross Sections and Photolysis Products. *J. Phys. Chem. A* **2008**, *112*, 9296−9303.

(19) Bertaux, J.-L.; Vandaele, A.-C.; Korablev, O.; Villard, E.; Fedorova, A.; Fussen, D.; Quémerais, E.; Belyaev, D.; Mahieux, A.; Montmessin, F.; et al. A Warm Layer in Venus' Cryosphere and High-Altitude Measurements of HF, HCl, H_2O and HDO. *Nature* **2007**, *450*, 646−649.

(20) Marcq, E.; Belyaev, D.; Montmessin, F.; Fedorova, A.; Bertaux, J. L.; Vandaele, A. C.; Neefs, E. An Investigation of the SO_2 Content of the Venusian Mesosphere Using SPICAV-UV in Nadir Mode. *Icarus* **2011**, *211*, 58−69.

(21) Young, A. T. The Clouds of Venus. *J. Atmos. Sci.* **1975**, *32*, 1125−1132.

(22) Delitsky, M. L.; Baines, K. H. Chemistry in the Venus Clouds: Sulfuric Acid Reactions and Freezing Behavior of Aqueous Liquid Droplets. *AAS/Division for Planetary Sciences Meeting Abstracts*; AAS: Washington, DC, 2015; p 217.02.

(23) Lane, J. R.; Kjaergaard, H. G. Fluorosulfonic Acid and Chlorosulfonic Acid: Possible Candidates for OH-Stretching Overtone-Induced Photodissociation. *J. Phys. Chem. A* **2007**, *111*, 9707−9713.

(24) Tamm, C. O. Acid Precipitation: Biological Effects in Soil and on Forest Vegetation. *Ambio* **1976**, *5*, 235−238.

(25) Ferenbaugh, R. W. Effects of Simulated Acid Rain on Phaseolus vulgaris L . (Fabaceae). *Am. J. Bot.* **1976**, *63*, 283−288.

(26) Havas, P. Injury to Pines in the Vicinity of a Chemical Processing Plant in Northern Finland. *Acta For. Fenn.* **1971**, 7555.

(27) Weinstein, L. Effects of Atmospheric Fluoride on Metabolic Constituents of Tomato and Bean Leaves. *Contrib. Boyce Thompson Inst.* **1961**, *21*, 215−231.

(28) Nutt, D. R.; Meuwly, M. Studying Reactive Processes with Classical Dynamics: Rebinding Dynamics in MbNO. *Biophys. J.* **2006**, *90*, 1191−1201.

(29) Danielsson, J.; Meuwly, M. Atomistic Simulation of Adiabatic Reactive Processes Based on Multi-State Potential Energy Surfaces. *J. Chem. Theory Comput.* **2008**, *4*, 1083.

(30) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, Revision D.01; Gaussian Inc.: Wallingford, CT, 2009.

(32) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *J. Comput. Chem.* **2011**, *32*, 2359−2368.

(33) Nelder, J.; Mead, R. A Simplex Method for Function Minimization. *Chem. Phys.* **1965**, *7*, 308−313.

(34) Yosa Reyes, J.; Brickel, S.; Unke, O. T.; Meuwly, M. HSO_3Cl: A Prototype Molecule for Studying OH-Stretching Overtone Induced Photodissociation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 6780−6788.

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

(35) Kottalam, J.; Case, D. A. Dynamics of Ligand Escape From the Heme Pocket of Myoglobin. *J. Am. Chem. Soc.* **1988**, *110*, 7690−7697.
(36) Souaille, M.; Roux, B. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.* **2001**, *135*, 40−57.
(37) Barker, J. R.; Nguyen, T. L.; Stanton, J. F.; Aieta, C.; Ceotto, M.; Gabas, F.; Kumar, T. J. D.; Li, C. G. L.; Lohr, L. L.; Maranzana, A.; et al. *MultiWell-2017 Software Suite*; University of Michigan: Ann Arbor, 1991.
(38) Schlegel, H. B.; Millam, J. M.; Iyengar, S. S.; Voth, G. A.; Daniels, A. D.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. *J. Chem. Phys.* **2001**, *114*, 9758−9763.
(39) Iyengar, S. S.; Schlegel, H. B.; Millam, J. M.; Voth, G. A.; Scuseria, G. E.; Frisch, M. J. Ab Initio Molecular Dynamics: Propagating the Density Matrix with Gaussian Orbitals. II. Generalizations Based on Mass-Weighting, Idempotency, Energy Conservation and Choice of Initial Conditions. *J. Chem. Phys.* **2001**, *115*, 10291−10302.
(40) Burkholder, J. B.; Mills, M.; McKeen, S. Upper Limit for the UV Absorption Cross Sections of H_2SO_4. *Geophys. Res. Lett.* **2000**, *27*, 2493−2496.
(41) Havey, D. K.; Feierabend, K. J.; Vaida, V. Ab Initio Study of H_2SO_4 Rotamers. *J. Mol. Struct.: THEOCHEM* **2004**, *680*, 243−247.
(42) Lane, J. R.; Kjaergaard, H. G.; Plath, K. L.; Vaida, V. Overtone Spectroscopy of Sulfonic Acid Derivatives. *J. Phys. Chem. A* **2007**, *111*, 5434−5440.
(43) Hintze, P. E.; Feierabend, K. J.; Havey, D. K.; Vaida, V. High-Resolution Spectroscopy of H_2SO_4, $HDSO_4$, and D_2SO_4 Vapor in the Region 1200−10,000 cm^{-1}. *Spectrochim. Acta, Part A* **2005**, *61*, 559−566.
(44) Feierabend, K. J.; Havey, D. K.; Brown, S. S.; Vaida, V. Experimental Absolute Intensities of the $4\nu_9$ and $5\nu_9$ O-H Stretching Overtones of H_2SO_4. *Chem. Phys. Lett.* **2006**, *420*, 438−442.
(45) Wrenn, S. J.; Butler, L. J.; Rowland, G. A.; Knox, C. J. H.; Phillips, L. F. The Necessity for Multiphoton Processes in the 193-nm. *J. Photochem. Photobiol., A* **1999**, *129*, 101−104.
(46) Robinson, T. W.; Schofield, D. P.; Kjaergaard, H. G. High Level Ab Initio Studies of the Excited States of Sulfuric Acid and Sulfur Trioxide. *J. Chem. Phys.* **2003**, *118*, 7226−7232.
(47) Yosa Reyes, J.; Nagy, T.; Meuwly, M. Competitive Reaction Pathways in Vibrationally Induced Photodissociation of H_2SO_4. *Phys. Chem. Chem. Phys.* **2014**, *16*, 18533−18544.
(48) Reiche, F.; Abel, B.; Beck, R.; Rizzo, T. Double-Resonance Overtone Photofragment Spectroscopy of trans-HONO. II. State- and Time-Resolved Dissociation and OH-Product State Distributions. *J. Chem. Phys.* **2002**, *116*, 10267−10276.
(49) Miller, Y.; Chaban, G. M.; Finlayson-Pitts, B. J.; Gerber, R. B. Photochemical Processes Induced by Vibrational Overtone Excitations: Dynamics Simulations for cis-HONO, trans-HONO, HNO_3, and HNO_3-H_2O. *J. Phys. Chem. A* **2006**, *110*, 5342−5354.
(50) Miller, Y.; Chaban, G. M.; Gerber, R. B. Theoretical Study of Anharmonic Vibrational Spectra of HNO_3, HNO_3-H_2O, HNO_4: Fundamental, Overtone and Combination Excitations. *Chem. Phys.* **2005**, *313*, 213−224.
(51) Fry, J.; Nizkorodov, S.; Okumura, M.; Roehl, C.; Francisco, J.; Wennberg, P. Cis-cis and trans-perp HOONO: Action Spectroscopy and Isomerization Kinetics. *J. Chem. Phys.* **2004**, *121*, 1432−1448.
(52) DeMore, W. B.; Sander, S. P.; Golden, D. M.; Hampson, R. F.; Kurylo, M. J.; Howard, C. J.; Ravishankara, A. R.; Kolb, C. E.; Molina, M. J. *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling*; JPL Publication 97-4; Jet Propulsion Lab: Pasadena, CA, 1997.
(53) Crim, F. F. Selective Excitation Studies of Unimolecular Reaction Dynamics. *Annu. Rev. Phys. Chem.* **1984**, *35*, 657−691.
(54) Kiefer, J. H.; Katopodis, C.; Santhanam, S.; Srinivasan, N. K.; Tranter, R. S. A Shock-Tube Laser-Schlieren Study of the Dissociation of 1,1,1-Trifluoroethane: An Intrinsic Non-RRKM Process. *J. Phys. Chem. A* **2004**, *108*, 2443−2450.
(55) Xu, H.; Kiefer, H.; Sivaramakrishnan, R.; Giri, B. R.; Tranter, R. S. Shock Tube Study of Dissociation and Relaxation in 1,1-Difluoroethane and Vinyl Fluorid. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4164−4176.
(56) Chackalackal, S. M.; Stafford, F. E. Infrared Spectra of Methane-, Fluoro-, and Chlorosulfonic Acids. *J. Am. Chem. Soc.* **1966**, *88*, 4815−4819.

DOI: 10.1021/acs.jpca.7b02950
J. Phys. Chem. A 2017, 121, 5079−5087

OH-Stretching Overtone Induced Dynamics in HSO_3F from Reactive Molecular Dynamics Simulations

Sebastian Brickel and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland

E-mail: m.meuwly@unibas.ch

Table S1: MP2 frequencies (in cm^{-1}) of HSO_3X (X = Cl, F) and the elimination TS (excluding the imaginary frequency) used for estimating the RRKM rates.

HSO_3F	HSO_3Cl	TS HSO_3F	TS HSO_3Cl
295	253	227	166
366	278	340	230
416	356	476	459
501	385	499	467
522	444	633	519
537	479	694	574
767	578	983	772
862	798	1002	1022
1190	1149	1268	1301
1233	1170	1450	1461
1483	1212	2035	1632
3795	3664		

*To whom correspondence should be addressed

4. Reactive Molecular Dynamics Simulations of the S_N2 Reaction

4.1. Introduction

The S_N2 reaction is an important bond forming/breaking bi-molecular reaction in chemical synthesis.[134] During the reaction the substituting nucleophile displaces a bound nucleophile from the molecule.[135] There have been numerous studies on the substitution reactions involving various different species.[80,83,136–144] Of particular interest is the influence of (polar) solvents on the reaction. Even though it is well established that in solution the activation energy is increased,[78–84] the effect of (polar) solvents on the reaction mechanism is still not well understood.

Three different reactive MD methods were applied to the reaction of $[Br\text{-}CH_3 + Cl]^-$: two methods combine empirical force fields for the reactants and for the products (MS-VALBOND and MS-ARMD), and the third method involves a neural network trained to reproduce thousands of reference energies.[145] Umbrella sampling simulations[101] were performed with all three methods to obtain the forward and reverse free energy barrier of the reaction. The two methods that are based on empirical force fields were further applied for simulations in aqueous solution without additional modifications for the change of environment. The goal of the work[24] was to compare the three methods, especially with regards to the preparation, transferability and the quality of the obtainable results.

4.2. Computational Methods

The S_N2 reaction of $[Br\text{-}CH_3 + Cl]^-$ to $[Br + H_3C\text{-}Cl]^-$ was studied using MS-ARMD and compared with two other reactive molecular dynamics simulation implementations (MS-VALBOND and

a neural network). The parametrisation of all three methods was performed with respect to MP2 (6-311++G(2d,2p) for MS-VALBOND and MS-ARMD and aug-cc-pVTZ for the neural network) level of theory using Gaussian09.[121] Furthermore, the MS-VALBOND parametrisation was also performed using CCSD/aug-cc-pVDZ as reference. This is possible since MS-VALBOND only requires a scan along the reaction coordinate for the fitting, therefore higher level of theory was amenable there. The correlation between these two quantum mechanical methods (MP2 and CCSD) was evaluated and found to be rather close to each other. Furthermore the energetics of both methods compared favorably with energies from the literature (the TS was found to be between 14.9 and 17.9 kcal/mol, the exothermicity of the reaction lies between 4.9 and 7.3 kcal/mol).[24,146]

Multi-State VALBOND describes the diagonal states, the two end points of the reaction with two Morse potentials, which were fitted to CCSD/aug-cc-pVDZ[110,147] and MP2/6-311++G(2d,2p) reference energies along the C-Br and C-Cl bonds in both molecules, respectively. The C-H bonds were described by a harmonic potential taken from CGenFF,[148] the H-C-H angles were described by a VALBOND term with sp^3 hybridization. The coupling between the reactant and product state was achieved by fitting scaled Morse potentials to the off-diagonal state using non-linear least square fitting (NL2SOL algorithm[149] in the R-program).[150] Details on the neural network are discussed elsewhere.[24,145]

4.2.1. MS-ARMD Parametrisation

An initial FF was obtained from SwissParam[127] for the reactant and the product molecules and anions. The Morse potential from the MS-VALBOND fit to CCSD/aug-cc-pVDZ was used to replace the harmonic bonds between C and Br and C and Cl in Br-CH_3 and Cl-CH_3, respectively. These two Morse potentials were fixed and not changed in the fitting. Furthermore, the harmonic bonds between C and H were replaced by Morse potentials. Both molecules were minimized, heated, and equilibrated. Snapshots extracted from a 250 ps MD simulation (NVE, $\Delta t = 0.25$ fs) calculated using the Leapfrog Verlet integrator[9] were used to calculate the MP2/6-311++G(2d,2p) reference data. For the fitting of the C-H bond and H-C-H angles, a downhill simplex algorithm[122] was utilized.

The so parameterized FF were then applied to the parametrisation of the charge transfer complexes ([Br-CH_3 + Cl]$^-$ and [Br + H_3C-Cl]$^-$). The fitting of the complexes was performed as for the molecules.

The Lennard Jones potential between the ion and the H atoms was replaced by generalized van der Waals parameters. The iterative fitting was completed with a total of 3291 structures. The connections between the reactant and product state was achieved by 3 pure Gaussian GAPOs, fitted to 21 IRC structures utilizing a GA. The quality of the fit is shown in Figure 4.1.

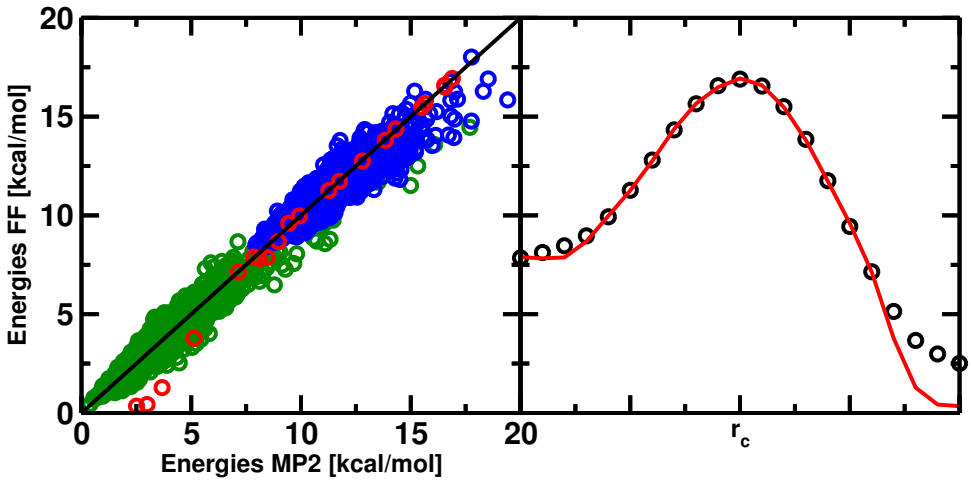


Figure 4.1.: Quality of parameterized MS-ARMD reactive FF compared to MP2/6-311++G(2d,2p) (left). [Br-CH_3 + Cl]$^-$: blue dots, [Br + H_3C-Cl]$^-$: green dots, IRC: red dots. Comparison of the energies between the IRC (right) from MP2/6-311++G(2d,2p) (black dots) with the parameterized MS-ARMD reactive FF (red line).

The reactant complex was solvated in a cubic periodic water box of size $30 \times 30 \times 30$ Å^3. The simulations in solution utilized the TIP3 water model[151] with the van der Waals parameters for the solute from the CGenFF force field.[148]

4.2.2. Free Energy Simulations

In order to obtain the free energy profile of the reaction, umbrella sampling simulations along the distance difference $r_c = d_{\text{C-Br}} - d_{\text{C-Cl}}$ were performed. In gas phase, as well as in aqueous solution, 50 equidistant umbrellas from $r_c = -2.5$ to $r_c = 2.5$ Å where calculated with 50 ps of dynamics each. The force constant was set to $k_{\text{umb}} = 50$ kcal/mol. MS-ARMD (and MS-VALBOND) MD simulations were performed using the CHARMM[9] program suite. The simulations were performed using the Velocity Verlet integrator ($\Delta t = 1$ fs). For the calculations in aqueous solution the system was minimized,

heated to 300 K and equilibrated for 1 ns. The potential of mean force was obtained via WHAM.[152]

Umbrella sampling with MS-ARMD yields a forward barrier of $\Delta G_f = 9.5$ kcal/mol (see Figure 4.2), compared to experimentally determined barrier of 9.5 $\pm$0.4 kcal/mol.[153,154] The reverse barrier is $\Delta G_r = 16.3$ kcal/mol, 0.8 kcal/mol below the experimental barrier.

Compared to the results obtained with the other two methods (see Table 4.1) the forward barrier is best described with MS-ARMD, while MS-VALBOND (both parametrisation) and the neural network overestimate the barrier by 1.7 - 3.5 kcal/mol. The reverse barrier is more closely reproduced, in comparison to experiment, by MS-VALBOND which yields 17.4 and 17.3 kcal/mol for the FF fitted to CCSD and MP2, respectively. The neural network overestimates the reverse barrier by 0.5 kcal/mol. It should be noted that the MS-ARMD results closely echo the energetic values for the IRC to which it was fitted. Hence the accurate reproduction of the forward barrier is most likely coincidental. The overall shape (see Figure 3 in Reference[24]) of the neural network and MS-ARMD agree well, while the MS-VALBOND line shapes are narrower around the TS and the increase of energy for larger ion-molecule separation is more pronounced in MS-VALBOND.

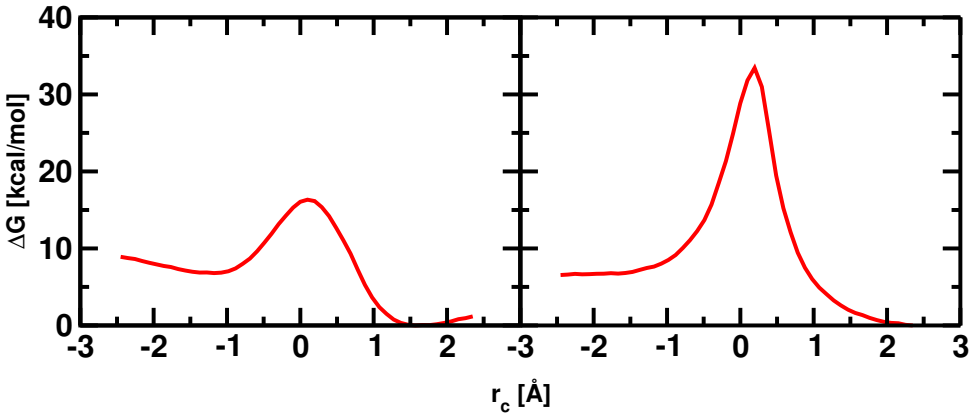


Figure 4.2.: Potential of mean force from umbrella sampling simulations in the gas phase (left) and in aqueous solution (right) using MS-ARMD. Forward reaction barrier corresponds to the barrier of the reaction [Br-CH_3 + Cl]$^-$ $\rightarrow$ [Br + H_3C-Cl]$^-$, reverse reaction barrier corresponds to the reverse reaction.

In order to obtain the PMF in solution it was necessary to increase the biasing potential strength

around the TS. The forward and reverse barrier are increased by 17.4 kcal/mol and 17.2 kcal/mol, respectively in comparison to gas phase. The two distinguishable minima in gas phase around the reactant and product conformation disappear in solution. Compared to experimental values the forward barrier is overestimated by 2.2 kcal/mol and the reverse barrier by 5.7 kcal/mol. This is in contrast to MS-VALBOND fitted to CCSD, which underestimates the forward barrier by 0.1 kcal/mol and the reverse barrier by 1.1 kcal/mol. A neural network was not trained for the reaction in solution since that would require a large amount of reference structures in order to capture the whole conformational space.

Table 4.1.: Comparison of forward (ΔG_f) and reverse (ΔG_r) barrier heights from the three utilized methods in gas phase (gas) and aqueous solution (sol), as well as experimental values from the literature. CCSD and MP2 stand for the MS-VALBOND parametrisation with respect to reference data calculated at the CCSD/aug-cc-pVDZ and MP2/6-311++G(2d,2p) level of theory, respectively. All values in kcal/mol. (Values taken from reference[24]).

	MS-VALBOND (CCSD/MP2)	MS-ARMD	neural network	Experiment
ΔG_f(gas)	13.0 / 11.2	9.5	12.1	9.5[153,154]
ΔG_r(gas)	17.4 / 17.3	16.3	17.6	17.1[155]
ΔG_f(sol)	24.6 / 21.8	26.9	–	24.7[156]
ΔG_r(sol)	26.7 / 28.2	33.5	–	27.8[157]

4.3. Discussion and Conclusions

The barrier heights in gas phase compare favorably for all three methods. A significant difference is that even when all three methods use a very similar level of theory for the reference data (all use MP2, but the neural network with a different basis set) the results still differ by more than 1 kcal/mol. The results obtained from condensed phase, where MS-ARMD and MS-VALBOND were used without further optimizations, show that the changes in MS-ARMD are stronger than for MS-VALBOND (both parametrisation). MS-ARMD is more sensitive to the changes of the environment. This becomes clearer when one compares the shape of the PMF. Where in the gas phase, MS-ARMD displays two minima, that disappear in solution. This is not the case for MS-VALBOND. The overestimation of the barrier height in aqueous solution from simulations with MS-ARMD was traced back to the lack

of a charge transfer function, which means that up to the TS the reactant charges were used. After passing the TS the product charges apply. In the present case, where the charge distribution changes dramatically during the reaction this procedure leads to a sharp peak at the TS. [24]

In terms of computational demands versus accuracy, it can be seen that MS-VALBOND only requires a few hundred reference structures and gives good results for both gas and condensed phase simulations. MS-ARMD needs more reference data, usually a few thousand, to achieve similar results. Both methods can be easily transferred from gas to condensed phase, even though the increased difference between both methods in condensed phase simulation points towards differences in sensitivity. MS-ARMD and MS-VALBOND are integrated in the CHARMM program suite. When taking the neural network as the most accurate description of the PES of the reaction, one can argue that the fully parameterized MS-ARMD FF yields an overall better description of the behaviour of the PES than the MS-VALBOND parametrisation in gas phase. This can be traced back to the origin of both methods. While MS-ARMD was designed with the description of gas phase reactions in mind, condensed phase reactions were the goal for MS-VALBOND.

5. Gas and Condensed Phase Simulations of the Claisen Rearrangement

The carbon-carbon bond forming[158] Claisen rearrangement[159] of allyl-vinyl-ether (AVE), AVE-2,6-dicarboxylate (AVE-$(CO_2)_2$), and the biochemically active molecule chorismate (see Figure 5.1) in gas and condensed phase, based on a well parameterized reactive FF of AVE, were studied using MS-ARMD. The focus was the analyzis of the mechanistic details, especially the effect of the catalysis in aqueous solution and in the enzyme chorismate mutase (CM). In comparison to the results in the previous chapter where the instantaneous switching of charges during the reaction lead to a spike in the PMF, this effect is small in the present case where the charge distribution of the reactant to product changes only slightly. In aqueous solution[85–87,160,161] and in the enzyme CM, which catalyses the Claisen rearrangement of chorismate to prephanate,[162] the reaction is speed-up by the stabilization of the cyclic TS. Compared to previous works by other groups, using QM/MM methods, MS-ARMD comes with approximately the cost of conventional force field, which allows extensive studies on the system. The reactive FF utilized here is parameterized to *ab initio* reference points at the MP2/6-311++G(2d,2p) level of theory. Such computational studies provide mechanistic insight into reaction mechanisms which are beyond current experiments.

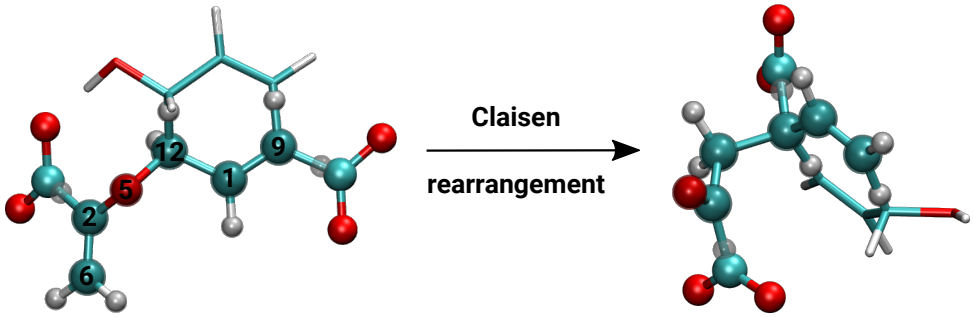


Figure 5.1.: Schematic representation of the three systems studies. The FF was fitted to AVE and pent-4-enal, shown as transparent ball and stick. Solid ball and stick shows AVE-$(CO_2)_2$ and its product, solid lines chorismate and prephanate. The important heavy atoms for the rearrangement are numbered on the reactant.

More O'Ferrall Jencks diagrams[163] projected on the CO (breaking bond) and CC (forming bond) distance reveal a step wise mechanism where bond breaking preceded bond formation, in line with existing experimental evidence. The analysis of the molecules positioning in the active site of the enzyme during the reaction shows entropic contribution to the enzymatic catalysis when comparing the three different molecules.

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

Cite This: J. Phys. Chem. B 2019, 123, 448−456

pubs.acs.org/JPCB

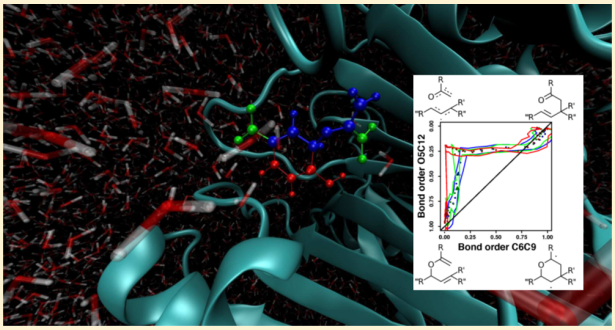
Molecular Determinants for Rate Acceleration in the Claisen Rearrangement Reaction

Sebastian Brickel and Markus Meuwly*

Department of Chemistry, University of Basel, Klingelbergstrasse 80, Basel CH-4056, Switzerland

Supporting Information

ABSTRACT: The Claisen rearrangement is a carbon−carbon bond-forming, pericyclic reaction of fundamental importance due to its relevance in synthetic and mechanistic investigations of organic and biological chemistry. Despite continued efforts, the molecular origins of the rate acceleration in going from the aqueous phase into the protein is still incompletely understood. In the present work, the rearrangement reactions for allyl-vinyl-ether (AVE), its dicarboxylated variant (AVE-$(CO_2)_2$), and the biologically relevant substrate chorismate are investigated in the gas phase, water, and in chorismate mutase. Only the rearrangement of chorismate in the enzyme shows a negative differential barrier when compared to the reaction in water, which leads to the experimentally observed catalytic effect for the enzyme. The molecular origin of this effect is the positioning of AVE-$(CO_2)_2$ and chorismate in the protein active site compared to AVE. Furthermore, in going from AVE-$(CO_2)_2$ to chorismate, entropic effects due to rigidification and ring formation are operative, which lead to changes in the rate. On the basis of "More O'Ferrall−Jencks" diagrams, it is confirmed that C−O bond breaking precedes C−C bond formation in all cases. This effect becomes more pronounced in going from the gas phase to the protein.



1. INTRODUCTION

The Claisen rearrangement, discovered at the beginning of the 20th century,[1] is an important [3,3]-sigmatropic rearrangement for high stereoselective[2] C−C bond formation.[3] Especially aliphatic Claisen rearrangement is of key importance in organic synthesis[4] and has been studied extensively by both experiment[5−9] and theory.[10−13]

The textbook example of a Claisen rearrangement is the reaction of allyl-vinyl-ether (AVE) to pent-4-enal[4] (see Figure 1A). In polar solvent, such as water, the stabilization of the transition state (TS) relative to the reaction in vacuum is the origin of the catalytic effect.[10,11,14] It was noted that the solvent effect is partly due to dipole−dipole interactions, which are more pronounced in polar solvents compared to nonpolar environments.[15,16] The reaction of chorismate to prephenate (see Figure 1C) in the shikimic acid pathway follows formally a Claisen rearrangement.[17] The shikimic acid pathway is operative in the biosynthesis of aromatic amino acids, for example, in bacteria, which establishes the biological relevance of the reaction. Therefore, this pathway is an attractive target for developing new antibiotics.[18] This has motivated numerous publications on enzymatic Claisen rearrangements in particular[18−28] and reactions with related substrates.[29−32] Compared to the reaction in aqueous solution, the enzymatic catalysis of the Claisen rearrangement reaction in chorismate mutase (CM) leads to a rate acceleration by ∼10^6 due to stabilization of the TS.[33] As a proxy for studying the interactions of

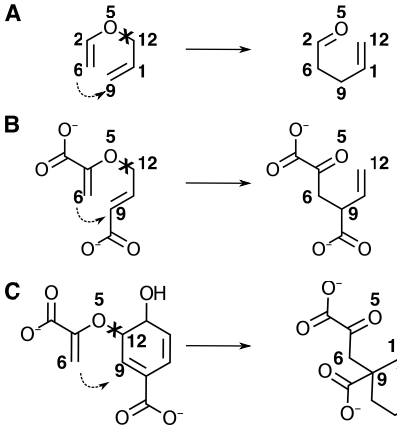


Figure 1. Claisen rearrangement reaction. A: allyl-vinyl ether → pent-4-enal, including numbering of the heavy atoms. B: AVE-2,6-dicarboxylate → pent-4-enal-dicarboxylate. C: chorismate → prephenate. "X" marks the O5−C12 bond that will be broken, and dashed arrow shows the two carbon atoms (C6, C9) that form a bond during rearrangement.

Received: November 14, 2018
Revised: December 6, 2018
Published: December 12, 2018

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

chorismate in CM, other related substrates (e.g., AVE-2,6-dicarboxylate (AVE-$(CO_2)_2$; see Figure 1B)) have also been used in the past.[27]

Despite the existing studies on Claisen rearrangement, molecular and certain mechanistic aspects of the reaction remain insufficiently characterized. One of them concerns the structural and molecular origins for the rate enhancement in going from the reaction in solution to that in the protein for chorismate mutase. Another one is the actual pathway between the reactant and product for which three possibilities are conceivable: (1) concerted (via an aromatic TS), (2) asynchronous via 1,4-diyl-like TS, or (3) asynchronous via a bis-allyl-like TS.[2] The clarification of these aspects—the molecular origins for the rate enhancement in different environments and the pathway between reactant and product—at the molecular level provides the motivation for the present study. Reactive atomistic simulations[34−36] are a meaningful approach for this as they can provide molecular-level insight into reactive processes on sufficiently long time scales in different chemical environments provided that a sufficiently accurate potential energy surface (PES) is available.

2. COMPUTATIONAL DETAILS

2.1. Molecular Dynamics (MD) Simulations. All simulations were performed with CHARMM,[37] including provision for forming and breaking bonds via MS-ARMD.[34] The simulations of AVE-$(CO_2)_2$ and chorismate used the parametrized reactive force field (FF) from AVE, except for the parameters directly related to the substitution (e.g., bond parameter from C2 to hydrogen in AVE, which is replaced by a C−C bond in AVE-$(CO_2)_2$ and chorismate), although further optimization as, for example, done for the dioxygenation reaction would be conceivable in principle.[36] This is possible since the substituted groups (e.g., CO_2) are not involved in the rearrangement in AVE-$(CO_2)_2$ and chorismate and remain largely unchanged during the reaction. This was confirmed by ab initio optimization of the reactant and product geometries.

Molecular dynamics simulations in the gas phase were performed at 300 K for a total of 100 ps (Δt = 0.25 fs) using the leap-frog Verlet integration scheme.[38] In the gas phase, the simulations of the three molecules (AVE, AVE-$(CO_2)_2$, and chorismate) started with the respective optimized reactant structure. The simulations in aqueous solution and in the enzyme were started after 50 steps of steepest descent and 50 steps of adopted basis Newton−Raphson, followed by 25 ps of NVT dynamics using the velocity Verlet integrator (Δt = 1 fs) followed by 500 ps (Δt = 2 fs) of free dynamics using SHAKE[39] to constrain all bonds involving hydrogen. Periodic boundary conditions together with the particle mesh Ewald[40] method was used for the long-range electrostatic interaction. The cutoffs for nonbonded, switching, and smoothing functions were 16, 14, and 12 Å, respectively.

For simulations in aqueous solution, the system was solvated in a pre-equilibrated TIP3P[41] water box (25 × 25 × 25 Å^3) obtained from CHARMM-GUI.[42] SHAKE[43] was used to constrain all bonds involving hydrogen atoms.

For the simulations in the enzyme, the PDB of the 2.2 Å X-ray structure of *Bacillus subtilis* CM (BsCM, PDB: 2CHT[44,45]) was used. BsCM is a homotrimer consisting of six chains, with the active sites located at the interface of two chains. The PDB of BsCM contains 1388 residues, 12 transition-state analogues (TSA,[46] 8-hydroxy-2-oxa-bicyclo[3.3.1]non-6-ene-3,5-dicarboxylic acid) of the inhibitor molecule, and 530 water molecules, which correspond to 4 trimers. One such trimer consisting of chains A to C, the crystallographic water, and the TSA was used as the starting point to set up the system. This subsystem consists of 6094 atoms, in total.

The simulations in the enzyme environment were set up by replacing the TSA of the inhibitor molecule between the two chains B and C in BsCM by AVE, AVE-$(CO_2)_2$, or chorismate. The remaining TSA molecules were removed. The enzyme, including the reactant molecule and the crystallographic water molecules, was solvated in a pre-equilibrated water box (including counter ions), of box size 116 Å × 116 Å × 116 Å. The overall system sizes were 149 218, 149 207, and 149 231 atoms, respectively.

2.2. Umbrella Sampling (US). Umbrella sampling (US)[47] was used to follow the reaction path of the Claisen rearrangement because the reaction barrier is high and not amenable to direct sampling. The reaction coordinate for the present application was defined as the difference between the breaking O−C bond and the forming C−C bond (see Figure 1), that is

$$r_c = d_{O5C12} - d_{C6C9} \quad (1)$$

This corresponds to the previously used reaction coordinate in QM/MM simulations on a chorismate TSA.[30] Simulations were carried out for equidistant windows from r_c = −2.6 to 2.2 Å (Δr_c = 0.2 Å), with k_{umb} = 200 kcal/mol/rad^2 for AVE and AVE-$(CO_2)_2$. For chorismate, the sampling was performed in equidistant steps in the range of r_c = −3.8 to 2.3 Å in order to capture the two minimum conformations calculated at the Møller−Plesset MP2/6-311++G(2d,2p) level of theory. Additional umbrellas were added where necessary.

For each window, statistics was accumulated after 5 ps of equilibration simulations. For the calculations in aqueous solution and BsCM, US simulations were performed starting each umbrella from a restart file produced in a previous US simulation with the same setting. This procedure prepares the system for the changes in the conformation of the molecules as well as reduces the difference from the initial condition to US conformation along the reaction coordinate. Finally, the analysis of the US simulations was carried out with the weighted histogram analysis method[48,49] with a tolerance of 0.01.

2.3. Parametrization of the Force Field. All reactive trajectories were generated with MS-ARMD[34] within the CHARMM[37] software package. MS-ARMD is a powerful, single-valued (by construction), reactive molecular dynamics implementation, which combines individual weighted surfaces $w_i(x) \times V_i(x)$ (i.e., connectivities), describing different states by parametrized FFs. The so-called GAPOs (Gaussian × polynomials) for describing the adiabatic barrier are added to form the global reactive potential energy surface (PES). GAPOs are calculated from the energy difference $\Delta V_{ij}(x)$ $(=V_j(x) - V_i(x))$ between two states.

Parametrized FFs for AVE and pent-4-enal were obtained by an iterative procedure, starting with reference parameters from SwissParam.[50] The energy at the equilibrium geometry of pent-4-enal was chosen as the global zero of energy. Using the initial force field, the equilibrium dynamics of AVE and pent-4-enal in the gas phase were followed for 250 ps at 300 K. A total of 3200 representative structures were extracted, and reference energies were determined at the MP2/6-311++G(2d,2p) level of theory using Gaussian09.[51] A downhill simplex algorithm[52] was used for fitting refined FF parameters. The bonds between

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

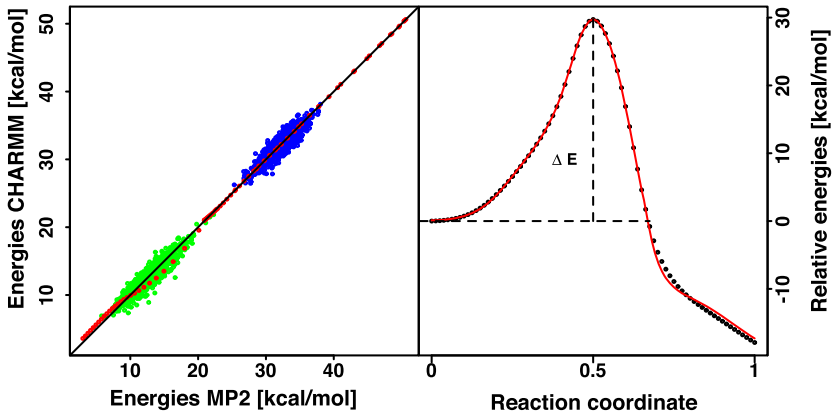


Figure 2. Left panel: comparison between the MP2/6-311++G(2d,2p) reference energies and the fitted FF. Energies for AVE (blue, RMSD = 0.82 kcal/mol), pent-4-enal (green, RMSD = 0.8 kcal/mol), and IRC (red, RMSD = 0.5 kcal/mol). Right panel: ab initio IRC (black points) and GAPO-fitted FF (red line). The right-hand-side *y*-axis shows the energy relative to the optimized reactant structure. The activation energy is $\Delta E^{\ddagger}$ = 29.7 kcal/mol.

O5−C12 in the reactant and C6−C9 in the product were described by a Morse oscillator. In a next step, these two FFs were combined to a reactive PES.

To parametrize the adiabatic barrier, the intrinsic reaction coordinate (IRC) of the reaction was calculated at the MP2/6-311++G(2d,2p) level of theory. Such calculations yield a barrier height of 29.7 kcal/mol, which agrees favorably with an activation energy of 30.6 kcal/mol determined from experiment[5] and values of >28 kcal/mol from further experiments.[3,9,53] Alternatively, density functional theory calculations at the BLYP/6-31G*[27] and B3LYP/6-31G*[10] levels yield somewhat lower barrier heights between 24.1 and 27.7 kcal/mol. Thus, the chosen level of theory was deemed appropriate. For the parametrization, 81 structures along the IRC were extracted and their MS-ARMD energy evaluated. A genetic algorithm for parametrizing the GAPOs was used to best reproduce the MP2 energies along the IRC path.

3. RESULTS AND DISCUSSION

3.1. Quality of the Force Field. All simulations are based on the parametrized force field for the conversion of AVE to pent-4-enal, which captures the energetics of the reactant and product to within 0.8 kcal/mol and the reaction path between the end points to within 0.5 kcal/mol; see Figure 2. The average root-mean-square deviation (RMSD) between the reference MP2 and the fitted FF energies is 0.7 kcal/mol, which establishes the quality of the FF. For further validation, the reactant and product geometries were minimized at the MP2 level and compared with the force field optimized structures (see Figure S1), and the normal mode frequencies were found to agree favorably, too (see Tables S1 and S2). The parametrized force field is provided in Tables S3−S10.

3.2. Free Energy Simulations for AVE. Using the parametrized, reactive force field, the barrier free energies $\Delta G^{\ddagger}$ for the rearrangement of AVE in the gas phase, aqueous solution, and the BsCM enzyme were determined using umbrella sampling (US) simulations. The activation free energies are $\Delta G^{\ddagger}$ = 23.4, 17.7, and 22.3 kcal/mol, respectively (see the top panel of Figure 3 and Table 1). Experimentally, barrier free energies have been determined in the gas phase and in apolar solvents to be ∼28 kcal/mol.[3,5,9] More recent selective integrate-over temperature simulations in water using density functional tight binding (DFTB) and the extended

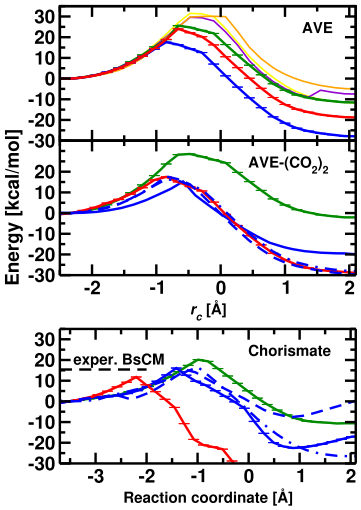


Figure 3. Potentials of mean force (PMF) for the Claisen rearrangement reactions for AVE (top), AVE-2,6-dicarboxylate (AVE-$(CO_2)_2$, middle), and chorismate (bottom) in the gas phase (green), aqueous solution (blue), and in the enzyme BsCM (red). Solid lines for forward simulations and dashed-dotted lines for the reverse reaction (product to reactant). Extended simulations (1 ns per window) are shown as dashed lines. The black dashed line indicates the experimentally measured value.[19] Error bars estimated from a bootstrap error analysis are overlaid by the line of the respective PMF and are of the size of the line width. DFTB3 forward simulations for AVE in the gas phase, water, and the enzyme are shown in yellow, violet, and orange, respectively.

simple point charge model water model yield $\Delta G^{\ddagger}$ = 22.6 kcal/mol, ∼5 kcal/mol higher than the present work. This compares with 29.6 kcal/mol obtained here from DFTB3[54] with the TIP3P water model. The activation enthalpy in the gas phase from DFTB calculations is 19.7 kcal/mol, which underestimates the experimentally determined $\Delta H^{\ddagger}$ (30.6 kcal/mol). On the other hand, the gas-phase free energy barrier height from US is $\Delta G^{\ddagger}$ = 31.8 kcal/mol, which corresponds to a TS stabilization of $\Delta\Delta G^{\ddagger}_{(\text{gas-solv})}$ = −2.2 kcal/mol.

The gas phase $\Delta G^{\ddagger}$ = 23.4 kcal/mol underestimates the barrier heights from the experiment by ∼5 kcal/mol. To better understand this difference, the distribution functions of all

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

Table 1. Free Energy Barrier (in kcal/mol) Calculated with MS-ARMD and US in Aqueous Solution and in the Enzyme *Bacillus subtilis* Chorismate Mutase (BsCM, 116 Å^3 Water Box)

	AVE	AVE-$(CO_2)_2$	chorismate	chorismate exper.
aqueous solution	17.7	15.2	16.8	24.5[a]
enzyme BsCM	22.3 (24.9)	18.1	12.4	15.4[b]

For AVE in the enzyme BsCM the result for the calculation in a 86 Å^3 water box is given in brackets. Experimental (exper.) $\Delta G^{\ddagger}$ from: [a]ref 33 (measured in Tris−HCl). [b]Values range from 15.0 to 16.2 depending on the protein.[19] All experimental $\Delta G^{\ddagger}$ values based on TST; analysis with an Arrhenius expression yields estimated values ~3 kcal/mol lower.[33]

bonds and valence angles involving non-hydrogen atoms were determined from the US simulations around the transition state. Compared with the MP2 structure of the TS (2.32 Å), the O5−C12 bond distribution peaks at 2.63 Å and for the C1−C12−O5 angle the MP2 value is 102° compared with the maximum of the distribution at 122°. All other bonds and angles agree to within 0.05 Å and 5°, respectively. Hence, the structure of the TS in the gas phase from US simulations is too open. Furthermore, it was found that despite the quite accurate parametrization of the reactant and product energies, the average harmonic frequencies from the force field differ by 89 and 120 cm^{-1} for the reactant and product, respectively. Therefore, two tests were carried out. First, the US simulations were repeated by increasing all product bond and angle force constants by 10%, and second, the US were rerun by placing a constraint (10.0 kcal/mol) on the O5−C12 bond to maintain it near the MP2 TS structure. Both modifications lead to an increase of the barrier, namely, by 5.3 kcal/mol when the force constants are changed and by 2.4 kcal/mol when running the simulations with the bias. Hence, for the gas-phase simulations, the force field could be further improved, and mild modifications in the force constants of the product already give near-quantitative agreement with experiment. However, the purpose of the present work is to investigate systematic aspects of the reaction for AVE, AVE-$(CO_2)_2$, and chorismate and to discuss them on a common footing and not to specifically optimize the FFs for individual ligands.

In solution, experiments were carried out in methanol−water mixtures (ranging from 100% MeOH to 25/75% MeOH/water), and the reaction barrier reduces by $\Delta\Delta G^{\ddagger}_{gas-sol}$ = −3.5 to −4.7 kcal/mol[55] with increasing water content. This compares with a reduction by $\Delta\Delta G^{\ddagger}_{gas-solv}$ = −6.2 kcal/mol from the present simulations in pure water and is expected due to stabilizing effects of the solvent environment. On the other hand, US simulations for this reaction in the protein active site yield $\Delta\Delta G^{\ddagger}_{water-prot}$ = +4.6 kcal/mol, that is, a slowdown of the reaction.

3.3. Free Energy Simulations for AVE-$(CO_2)_2$. Next, the rearrangement reaction of AVE-2,6-dicarboxylate to pent-4-enal-dicarboxylate was investigated in the three environments. The gas-phase PMF (green line in the middle panel of Figure 3) for AVE-$(CO_2)_2$ shows an increase in barrier height compared to AVE (32.2 vs 23.4 kcal/mol). This can be explained by the destabilizing effect of the two negatively charged carboxylate groups, which are insufficiently screened in the gas phase. In aqueous solution, the barrier is

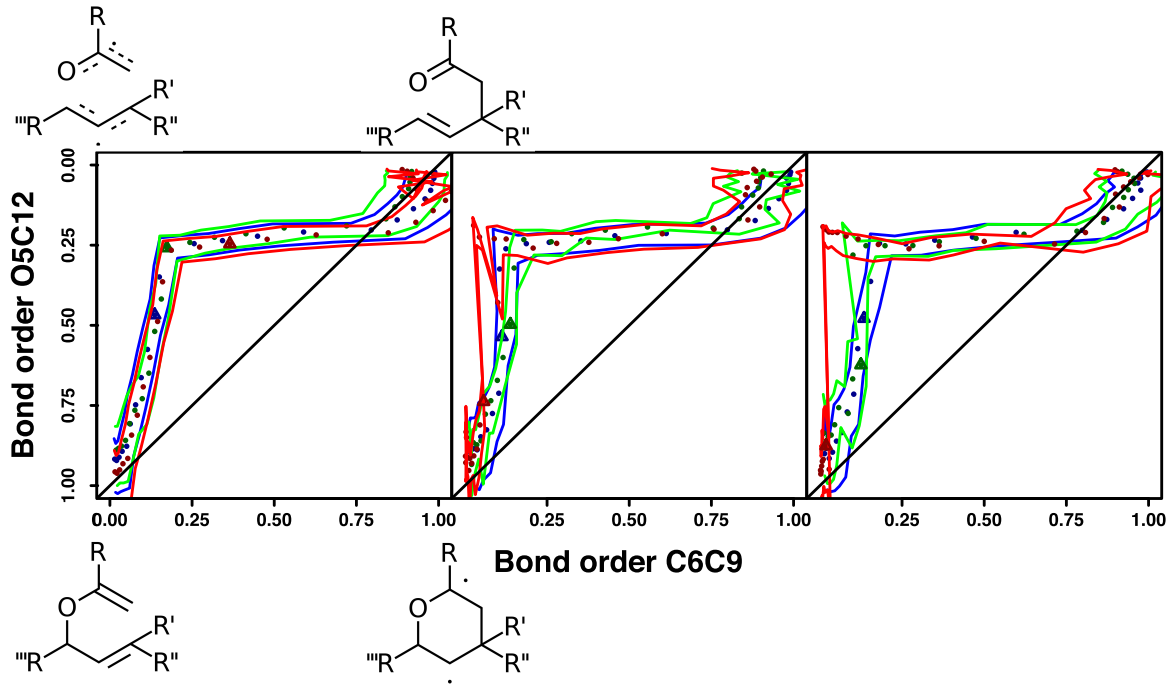


Figure 4. More O'Ferrall−Jencks diagrams for Claisen rearrangement in the gas phase (left), aqueous solution (middle), and in the enzyme BsCM (right) for AVE (blue), AVE-2,6-dicarboxylate (green), and chorismate (red). Bond orders are defined using the interatomic distance for the bonds from the respective ab initio optimized structure in the Pauling relations (eq 2). From secondary tritium isotope effect in solution, the TS of chorismate has been localized at a bond order of ≈0.6 for the O5−C12 separation with no detectable C6−C9 bond.[65] For MOFJ diagrams by molecule, see Figure S3.

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

considerably reduced (by $\Delta\Delta G^{\ddagger}_{gas-solv}$ = 17 kcal/mol), as expected. The TS stabilization for AVE-$(CO_2)_2$ in solution is enhanced compared to AVE due to screening of the repelling charges of the two carboxylate groups during the rearrangement. Extended simulations (1 ns per umbrella; see dashed trace in the middle panel of Figure 3) show a change in the free energy of the product as the surrounding water adapts to the structural changes of the ligand. Such changes on the product side due to slow reorganization have also been found in ligand binding simulations for cyano-benzene in lysozyme.[56] However, for the present work, the forward barrier between the reactant and the TS is of primary concern (and can be directly compared with experiment). The forward barrier for the extended sampling increases by 2.5 kcal/mol. This was confirmed by US simulations starting from the solvated product, rather than the solvated reactant, which gives a similar result as the 1 ns forward US simulations (dashed-dotted trace, middle panel Figure 3) and suggests that the simulations are close to converged. No experimental data for the dicarboxylate is available, but reactions for the monocarboxylate in Me_2SO/H_2O (ratio 9:1) mixtures have been reported,[8] which, however, are not ideal for direct comparison with the computations. For the reaction in the protein, the computed $\Delta\Delta G^{\ddagger}$ increases to 18.1 kcal/mol, which indicates that the reaction actually slows down in the protein.

3.4. Free Energy Simulations for Chorismate. Finally, the Claisen rearrangement was investigated using the biologically relevant substrate, chorismate (see the bottom panel in Figure 3). In addition to the two $-CO_2^-$ groups, the allyl is completed to form a six-membered ring; see Figure 1C. This makes the substrate in the reactant state less flexible and bulkier compared to the AVE-2,6-dicarboxylate. Contrary to the two other systems, the activation free energy monotonically decreases from $\Delta G^{\ddagger}_{gas}$ = 20.5 kcal/mol in the gas phase to $\Delta G^{\ddagger}_{prot}$ = 12.4 kcal/mol in the protein. In aqueous solution, the free energy barrier is intermediate (16.8 kcal/mol) to the two and differs from experiment (24.5 kcal/mol, carried out in 0.02 M Tris−HCl buffer) by 7.7 kcal/mol. However, the results are in quite good agreement with QM/MM studies (B3LYP/6-31G(d)/CHARMM27), which yielded 17.4 ± 1.9 kcal/mol.[30] Using an empirical valence bond (EVB) model parametrized to reproduce the barrier in water, the free energy barrier height in the solvent was found to be 25.1 kcal/mol.[57] As for the results of AVE-$(CO_2)_2$ in solution, extending the sampling in the US simulations to 1 ns for every window increases the free energy of the product by around 15 kcal/mol (dashed blue vs solid blue line in the bottom panel of Figure 3), showing the influence of environmental adaptation on the PMF.[56,58] The barrier height for the forward reaction is not significantly (decrease by 0.8 kcal/mol) influenced.

3.5. Mechanistic Details. To clarify the mechanistic details involved in the rearrangement, "More O'Ferrall−Jencks" (MOFJ) diagrams[59,60] were generated for all US simulations (50 ps per window). Such diagrams involve two progression coordinates and are useful to illustrate their coupling during the course of a reaction. The corners of the diagram correspond to the reactant, product, and the two nonconcerted extremes; see the left panel in Figure 4. MOFJ diagrams are usually generated from coordinates that quantify the existence of a bond, such as a distance, bond order, or degree of bond making/breaking. In the present case (see Figure 4), the arithmetic average (dots) and the Kernel density estimate (KDE, using the R-program)[61,62] minimum and maximum (lines) distribution of the umbrella trajectories are reported. For computing the bond order n', the Pauling relation

$$n' = \exp\left(\frac{D_1 - D(n')}{C}\right) \quad (2)$$

is used, where D_1 is the reference interatomic distance from the ab initio optimized structures, $D(n')$ is the instantaneous separation during the simulations, and C = 0.6 Å was used[13,63,64] (for the dependence on the choice of C, see Figure S2).

Figure 4 shows that the rearrangement reaction in all three environments follows a stepwise mechanism, in which first the O5−C12 bond is partially broken and leads toward a bis-allyl-like structure (top-left corner), followed by formation of the C6−C9 bond. Running the reaction in polar solvent and in the protein leads to steeper curves during the first phase of the reaction; see Figure 4 (middle panel). In all cases, the transition states (indicated by solid triangles) are early, that is, they occur during the C−O bond-breaking process. This agrees with kinetic isotope experiments, according to which C−O cleavage always precedes C−C bond formation.[66] For the nonenzymatic thermal rearrangement of chorismate to prephenate, the measured kinetic isotope effects[65,66] indicate that at the TS the C−O bond is about 40% broken, but little or no C−C bond is formed, consistent with the MOFJ diagrams assuming that bond order and C−O distance are linearly related.

Unbiased MD simulations starting at the TS preserve the overall structure of the MOFJ diagrams whereby the TS is shifted toward a more "broken" C−O bond; see Figure 5. These unbiased simulations further establish that C−O bond

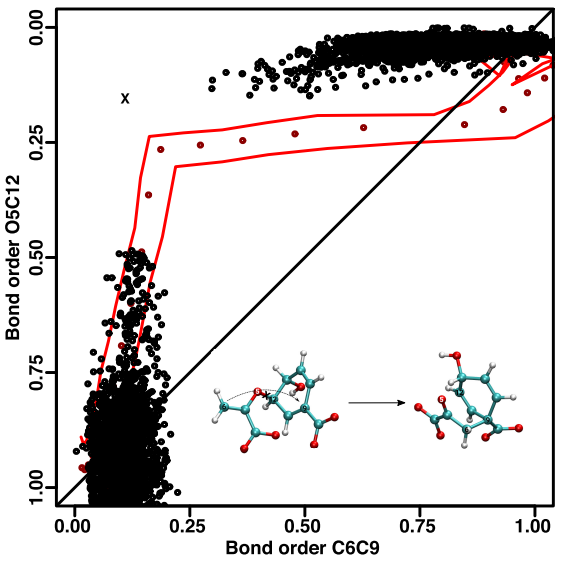


Figure 5. More O'Ferrall−Jencks diagrams for Claisen rearrangement of chorismate in the gas phase. The red points are the maxima of each distribution from the US simulations and the solid red line reports the minima and maxima of the corresponding KDE. Bond orders are defined using the interatomic distance for the bonds from the respective ab initio optimized structure in the Pauling relation (eq 2). Black points show the distribution of 32 free dynamics started from TS structures extracted from the US. The black X marks the center of gravity of the distribution around the TS.

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

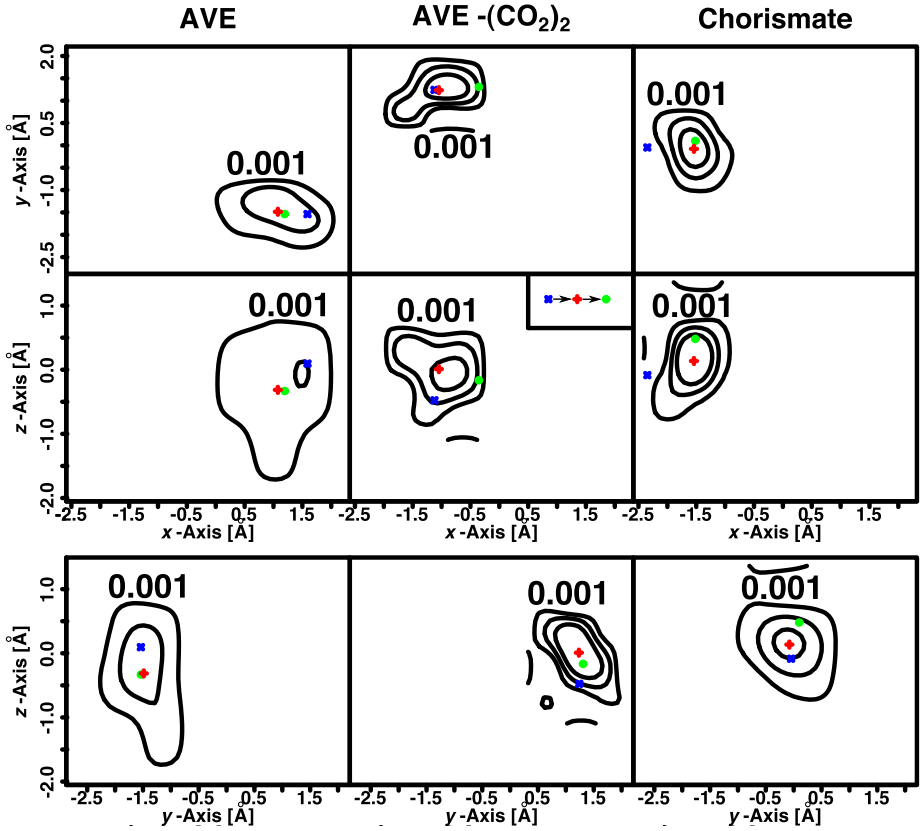


Figure 6. Normalized probability distributions of the *x*-, *y*-, and *z*-coordinates of the center of mass of the heavy atom framework of AVE (left), AVE-$(CO_2)_2$ (middle), and chorismate (right) at the TS in the enzyme BsCM. The centers of gravity of the reactant (blue x), TS (red plus), and product (green dot) are displayed in each panel. The origin of the coordinate system is the position of the transition-state analogue in the X-ray structure (see Figure 7). The enzyme backbone of the US trajectories was aligned to the enzyme backbone of the X-ray structure. For the distribution for the reactant and the product states, see Figures S4 and S5.

breaking precedes C−C formation. The average O5−C12 and C6−C9 bond lengths at the transition state from the US simulations are 2.0 and 2.6 Å, respectively. At the B3LYP/6-311++G(2d,2p) level, the energy of this configuration is ≈40% higher than the minimum energy for the constrained C−C bond at the TS value, which also supports the notion that the C−O bond is appreciably weakened, whereas the C−C bond is still not formed. In the enzyme, the measured[65] isotope effect of 0.99 suggests an even earlier transition state than in water, which is consistent with the free energy simulations (see Figure 3).

For AVE-$(CO_2)_2$ and especially chorismate, the effect of TS stabilization is further enhanced in the enzyme BsCM due to spatial constraints. This can be seen by the pronounced lag in C6−C9 bond formation of these two substrates in BsCM. AVE in BsCM shows a distribution intermediate to that in the gas phase and aqueous solution. This difference between AVE and the other two substrates is indicative of the importance of the carboxylate groups in AVE-$(CO_2)_2$ and chorismate for the catalytic step in BsCM in locking the substrate into place. The importance of the carboxylate groups was also shown by EVB simulations[57] and by measurements of rate constants of different chorismate-like substrates.[67] However, the barrier heights for AVE and AVE-$(CO_2)_2$ in the enzyme compared to simulations in aqueous solution increase. This anticatalytic effect decreases from AVE to AVE-$(CO_2)_2$ due to the presence of the $-CO_2$ groups and may reduce to close to zero when sampling is extended in the US (see above: increase of the free energy barrier height in water by 2.5 kcal/mol). Additional entropic effects lead to the catalytic effect found for chorismate.

Starting from 32 chorismate structures extracted from the US simulations at the TS in the gas phase, free MD simulations were performed. The resulting MOFJ diagrams from the unbiased simulations are consistent with those from the US simulation (see Figure 5) with the center of gravity from the unbiased simulations for the TS structure somewhat more advanced along the O5−C12 bond cleavage (see black X in Figure 5) compared to that from US.

The differences in the enzymatic activity for the three substrates can also be related to spatial determinants within the active site. For this, the position of the center of mass (CoM) of the molecules within the active site is analyzed. Figure 6 shows the normalized probability distributions of *x*-, *y*-, and *z*-coordinates of the CoM of the heavy atom framework common to all three ligands from the umbrella sampling simulations at the TS of the PMF. A broad or narrow distribution is indicative of a "loose" or "tight" TS, respectively.

With respect to the localization within the active site, AVE-$(CO_2)_2$ and chorismate occupy comparable regions, whereas AVE does not. The TS for AVE is displaced by 2 and 3 Å along the *x*- and *y*-axis, respectively. Furthermore, the widths of the distributions decrease in going from AVE, over AVE-$(CO_2)_2$ to chorismate, that is, the distributions change from loose to tight,

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

as does the TS itself; see Figure 4. Specifically, the widths along the *x*- and *z*-directions are different for AVE on one side and AVE-$(CO_2)_2$ and chorismate on the other. Owing to the presence of the CO_2^- groups, the positioning of the CoM of AVE-2,6-dicarboxylate and chorismate is similar to one another. However, AVE-$(CO_2)_2$ shows a broader distribution in the *x*-direction and a repositioning along *y*, indicating increased flexibility of the CoM compared to chorismate.

Figure 6 also shows the centers of gravity of the reactant (blue cross), the TS (red plus), and the product (green dot) of the three ligands in the active site. For AVE, the center of gravity of the TS overlays with that of the product, which is not the case for the other two molecules. The distributions of the centers of gravities of AVE-$(CO_2)_2$ show movement of the ligand during the reaction, except for the *y*-axis position. For chorismate, the center of gravity is further displaced between the reactant and the TS compared to AVE-$(CO_2)_2$, whereas the movement between TS and product is reduced. This suggests that stabilization of the transition state is coupled to displacement in configurational space and is indicative of TS stabilization rather than substrate binding. (Figure 7)

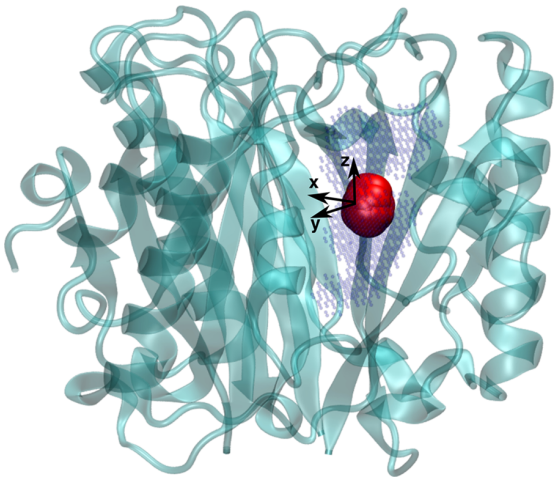


Figure 7. Overview of the simulation system for chorismate mutase and the center of mass of the heavy atom framework of chorismate (red) from the umbrella at the TS in Figure 3. Blue spheres delineate the active site from pocket mapping.[68,69] The origin of the coordinate system for the analysis of the localization for each ligand with the three Cartesian axes is shown in black.

4. CONCLUSIONS

The barrier height for the reaction of chorismate in BsCM is 12.4 kcal/mol, which yields a speedup by 3 orders of magnitude as estimated from the transition-state theory. Hence, in going from AVE and AVE-$(CO_2)_2$ to chorismate and using the same reactive PES, the simulations find that for smaller substrates, the rate slows down when going from water to the protein as the environment, whereas for the largest substrate (chorismate) they correctly find that the protein accelerates the reaction. Considering the changes of +4.6 kcal/mol (AVE), +2.9 kcal/mol (AVE-$(CO_2)_2$), and −4.4 kcal/mol (chorismate) in the activation free energies and correlating them with the actual chemical modifications suggest that both electrostatic stabilization (AVE → AVE-$(CO_2)_2$) and entropic contributions (AVE-$(CO_2)_2$ → chorismate, through the rigidification and larger size of chorismate) lead to the rate enhancement observed for chorismate in CM.

The experimentally measured acceleration is 6 orders of magnitude,[19] which is underestimated by the present simulations. However, as already mentioned, quantitative agreement for individual systems in specific environments could be further improved by a dedicated force field parametrization compared to the experiment.[57] The present approach, on the contrary, focuses on a systematic assessment of the reaction for the three different systems without specific optimization of the energy function in a particular environment.

In summary, the Claisen rearrangement reactions of AVE, AVE-$(CO_2)_2$, and chorismate in the gas phase, aqueous solution, and BsCM follow a stepwise mechanism in which the O5−C12 bond breaks prior to C6−C9 bond formation, especially in aqueous solution and in BsCM. The analysis of the TS position in the active site of BsCM reveals that the lack of catalytic effect on AVE is due to its loose positioning, insufficient interaction with, and TS stabilization by the active site of the enzyme. Major contributions to localizing the substrate in the active site of BsCM originate from the CO_2^- groups. This, together with the probability distributions in the reactant, TS, and product states, suggests that entropic factors must also be considered when interpreting differences between the systems, specifically (but not only) in the protein environment. This requires extended sampling of the configurational space, which is currently only possible with reactive force fields and outside the scope of mixed QM/MM MD simulations. The present work provides molecular-level insight into the Claisen rearrangement of chorismate in BsCM that will serve as a proxy for using computational methods for future development in protein engineering.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.8b11059.

Computational details and validation of the force fields together with all force field parameters for the ligands, additional More O'Ferrall−Jencks analysis by systems and center of mass distributions for reactants and products in the active site of BsCM (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: m.meuwly@unibas.ch. Phone: +41 (0)61 267 38 21. Fax: +41 (0)61 207 38 55.

ORCID

Sebastian Brickel: 0000-0003-1972-0393

Markus Meuwly: 0000-0001-7930-8806

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Prof. Adrian J. Mulholland, Donald Hilvert, and Dan Singleton for insightful comments. The authors gratefully acknowledge the financial support from the Swiss National Science Foundation through the NCCR MUST and grant 200021-117810.

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

REFERENCES

(1) Claisen, L. Über Umlagerung von Phenol-allyläthern in C-Allyl-phenole. *Chem. Ber.* **1912**, *45*, 3157−3166.
(2) Iwakura, I. The Experimental Visualisation of Molecular Structural Changes During Both Photochemical and Thermal Reactions by Real-Time Vibrational Spectroscopy. *Phys. Chem. Chem. Phys.* **2011**, *13*, 5546−5555.
(3) Coates, R. M.; Rogers, B. D.; Hobbs, S. J.; Curran, D. P.; Peck, D. R. Synthesis and Claisen Rearrangement of Alkoxyallyl Enol Ethers. Evidence for a Dipolar Transition State. *J. Am. Chem. Soc.* **1987**, *109*, 1160−1170.
(4) Ziegler, F. E. The Thermal, Aliphatic Claisen Rearrangement. *Chem. Rev.* **1988**, *88*, 1423−1452.
(5) Schuler, F. W.; Murphy, G. W. The Kinetics of the Rearrangement of Vinyl Allyl Ether. *J. Am. Chem. Soc.* **1950**, *72*, 3155−3159.
(6) Gajewski, J.; Brichford, N. Secondary Deuterium Kinetic Isotope Effects in the Aqueous Claisen Rearrangement: Evidence Against an Ionic Transition State. *J. Am. Chem. Soc.* **1994**, *116*, 3165−3166.
(7) Gajewski, J.; Conrad, N. Variable Transition State Structure in 3,3-Sigmatropic Shifts from α-Secondary Deuterium Isotope Effects. *J. Am. Chem. Soc.* **1979**, *101*, 6693−6704.
(8) Gajewski, J. J.; Jurayj, J.; Kimbrough, D. R.; Gande, M. E.; Ganem, B.; Carpenter, B. K. The Mechanism of Rearrangement of Chorismic Acid and Related Compounds. *J. Am. Chem. Soc.* **1987**, *109*, 1170−1186.
(9) Burrows, C.; Carpenter, B. Substituent Effects on the Aliphatic Claisen Rearrangement. 1. Synthesis and Rearrangement of Cyano-Substituted Allyl Vinyl Ethers. *J. Am. Chem. Soc.* **1981**, *103*, 6983−6984.
(10) Guest, J. M.; Craw, J. S.; Vincent, M. A.; Hillier, I. H. The Effect of Water on the Claisen Rearrangement of Allyl Vinyl Ether: Theoretical Methods Including Explicit Solvent and Electron Correlation. *J. Chem. Soc., Perkin Trans. 2* **1997**, 71−74.
(11) Severance, D. L.; Jorgensen, W. L. Effects of Hydration on the Claisen Rearrangement of Allyl Vinyl Ether from Computer Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10966−10968.
(12) Davidson, M. M.; Hillier, I. H.; Vincent, M. A. The Claisen Rearrangement of Allyl Vinyl Ether in the Gas Phase and Aqueous Solution. Structures and Energies Predicted by High-Level Ab Initio Calculations. *Chem. Phys. Lett.* **1995**, *246*, 536−540.
(13) Meyer, M. P.; DelMonte, A. J.; Singleton, D. A. Reinvestigation of the Isotope Effects for the Claisen and Aromatic Claisen Rearrangements: The Nature of the Claisen Transition States. *J. Am. Chem. Soc.* **1999**, *121*, 10865−10874.
(14) Cramer, C. J.; Truhlar, D. G. What Causes Aqueous Acceleration of the Claisen Rearrangement? *J. Am. Chem. Soc.* **1992**, *114*, 8794−8799.
(15) White, W. N.; Wolfarth, E. F. The o-Claisen rearrangement. VIII. Solvent effects. *J. Org. Chem.* **1970**, *35*, 2196−2199.
(16) Acevedo, O.; Armacost, K. Claisen Rearrangements: Insight into Solvent Effects and "on Water" Reactivity from QM/MM Simulations. *J. Am. Chem. Soc.* **2010**, *132*, 1966−1975.
(17) Ganem, B. From Glucose to Aromatics: Recent Developments in Natural Products of the Shikimic Acid Pathway. *Tetrahedron* **1978**, *34*, 3353−3383.
(18) Davidson, M. M.; Guest, J. M.; Craw, J. S.; Hillier, I. H.; Vincent, M. A. Conformational and Solvation Aspects of the Chorismate - Prephenate Rearrangement Studied by Ab Initio Electronic Structure and Simulation Methods. *J. Chem. Soc., Perkin Trans. 2* **1997**, 1395−1400.
(19) Kast, P.; Asif-Ullah, M.; Hilvert, D. Is Chorismate Mutase a Prototypic Entropy Trap? - Activation Parameters for the Bacillus Subtilis Enzyme. *Tetrahedron Lett.* **1996**, *37*, 2691−2694.
(20) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. Insights Into Enzyme Catalysis from QM/MM Modelling: Transition State Stabilization in Chorismate Mutase. *Mol. Phys.* **2003**, *101*, 2695−2714.
(21) Lever, G.; Cole, D. J.; Lonsdale, R.; Ranaghan, K. E.; Wales, D. J.; Mulholland, A. J.; Skylaris, C. K.; Payne, M. C. Large-Scale Density Functional Theory Transition State Searching in Enzymes. *J. Phys. Chem. Lett.* **2014**, *5*, 3614−3619.
(22) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J. Theoretical QM/MM Studies of Enzymatic Pericyclic Reactions. *Interdiscip. Sci.: Comput. Life Sci.* **2010**, *2*, 115−131.
(23) Ferrer, S.; Martí, S.; Andrés, J.; Moliner, V.; Tuñón, I.; Bertrán, J. Molecular Mechanism of Chorismate Mutase Activity of Promiscuos MbtI. *Theor. Chem. Acc.* **2011**, *128*, 601−607.
(24) Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J.; Field, M. J. A Hybrid Potential Reaction Path and Free Energy Study of the Chorismate Mutase Reaction. *J. Am. Chem. Soc.* **2001**, *123*, 1709−1712.
(25) Roca, M.; Vardi-Kilshtain, A.; Warshel, A. Toward Accurate Screening in Computer-Aided Enzyme Design. *Biochemistry* **2009**, *48*, 3046−3056.
(26) Madurga, S.; Vilaseca, E. SCRF Study of the Conformational Equilibrium of Chorismate in Water. *Phys. Chem. Chem. Phys.* **2001**, *3*, 3548−3554.
(27) Wiest, O.; Houk, K. N. Stabilization of the Transition State of the Chorismate-Prephenate Rearrangement: An ab Initio Study of Enzyme and Antibody Catalysis. *J. Am. Chem. Soc.* **1995**, *117*, 11628−11639.
(28) Hur, S.; Bruice, T. C. The Mechanism of Catalysis of the Chorismate to Prephenate Reaction by the *Escherichia Coli* Mutase Enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1176−1181.
(29) Vance, R.; Rondan, N.; et al. Transition Structures for the Claisen Rearrangement. *J. Am. Chem. Soc.* **1988**, *110*, 2314−2315.
(30) Claeyssens, F.; Ranaghan, K. E.; Lawan, N.; Macrae, S. J.; Manby, F. R.; Harvey, J. N.; Mulholland, A. J. Analysis of Chorismate Mutase Catalysis by QM/MM Modelling of Enzyme-Catalysed and Uncatalysed Reactions. *Org. Biomol. Chem.* **2011**, *9*, 1578−1590.
(31) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. Transition State Stabilization and Substrate Strain in Enzyme Catalysis: Ab Initio QM/MM Modelling of the Chorismate Mutase Reaction. *Org. Biomol. Chem.* **2004**, *2*, 968−980.
(32) Carlson, H. A.; Jorgensen, W. L. Monte Carlo Investigations of Solvent Effects on the Chorismate to Prephenate Rearrangement. *J. Am. Chem. Soc.* **1996**, *118*, 8475−8484.
(33) Andrews, P. R.; Smith, G. D.; Young, I. G. Transition-State Stabilization and Enzymic Catalysis. Kinetic and Molecular Orbital Studies of the Rearrangement of Chorismate to Prephenate. *Biochemistry* **1973**, *12*, 3492−3498.
(34) Nagy, T.; Reyes, J. Y.; Meuwly, M. Multisurface Adiabatic Reactive Molecular Dynamics. *J. Chem. Theory Comput.* **2014**, *10*, 1366−1375.
(35) Soloviov, M.; Das, A. K.; Meuwly, M. Structural Interpretation of Metastable States in Myoglobin-NO. *Angew. Chem., Int. Ed.* **2016**, *55*, 10126−10130.
(36) Das, A. K.; Meuwly, M. Kinetics and Structural Interpretation of Competitive Ligand Binding for NO Dioxygenation in Truncated Hemoglobin N. *Angew. Chem., Int. Ed.* **2018**, *57*, 3509−3513.
(37) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(38) Verlet, L. Computer Experiments on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98−103.
(39) VanGunsteren, W.; Berendsen, H. Algorithms for Macro-molecular Dynamics and Constraint Dynamics. *Mol. Phys.* **1977**, *34*, 1311−1327.
(40) Essmann, U.; Perera, L.; Darden, M. L. B.; Lee, H.; Pedersen, L. G.; Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; et al. A Smooth Particle Mesh Ewald Method A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

(41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(42) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859−1865.
(43) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(44) Chook, Y. M.; Ke, H.; Lipscomb, W. Crystal Structures of the Monofunctional Chorismate Mutase from *Bacillus subtilis* and its Complex with a Transition State Analog. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8600−8603.
(45) Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; Shindyalov, I.; Bourne, P. The Protein Data. www.rcsb.org, 2000.
(46) Bartlett, P. A.; Johnson, C. R. An Inhibitor of Chorismate Mutase Resembling the Transition-State Conformation. *J. Am. Chem. Soc.* **1985**, *107*, 7792−7793.
(47) Kottalam, J.; Case, D. A. Dynamics of Ligand Escape from the Heme Pocket of Myoglobin. *J. Am. Chem. Soc.* **1988**, *110*, 7690−7697.
(48) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(49) Souaille, M.; Roux, B. Extension to the Weighted Histogram Analysis Method: Combining Umbrella Sampling with Free Energy Calculations. *Comput. Phys. Commun.* **2001**, *135*, 40−57.
(50) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: A Fast Force Field Generation Tool for Small Organic Molecules. *J. Comput. Chem.* **2011**, *32*, 2359−2368.
(51) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.;et al. *Gaussian Inc 09*, revision D.01; Gaussian Inc.: Wallingford, CT, 2009.
(52) Nelder, J.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308−313.
(53) Zhang, J.; Yang, Y. I.; Yang, L.; Gao, Y. Q. Dynamics and Kinetics Study of In-Water Chemical Reactions by Enhanced Sampling of Reactive Trajectories. *J. Phys. Chem. B* **2015**, *119*, 14505−14514.
(54) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(55) Brandes, E.; Grieco, P.; Gajewski, J. Effect of Polar-Solvents on the Rates of Claisen Rearrangements - Assessment of Ionic Character. *J. Org. Chem.* **1989**, *54*, 515−516.
(56) Mondal, P.; Meuwly, M. Vibrational Stark Spectroscopy for Assessing Ligand-Binding Strengths in a Protein. *Phys. Chem. Chem. Phys.* **2017**, *19*, 16131−16143.
(57) Štrajbl, M.; Shurki, A.; Kato, M.; Warshel, A. Apparent NAC Effect in Chorismate Mutase Reflects Electrostatic Transition State Stabilization. *J. Am. Chem. Soc.* **2003**, *125*, 10228−10237.
(58) El Hage, K.; Mondal, P.; Meuwly, M. Free Energy Simulations for Protein Ligand Binding and Stability. *Mol. Simul.* **2018**, *44*, 1044−1061.
(59) O'Ferrall, R. A. M. Relationships between E2 and E1cB Mechanisms of β-Elimination. *J. Chem. Soc.* **1970**, 274−277.
(60) Jencks, W. P. General Acid-Base Catalysis of Complex Reactions in Water. *Chem. Rev.* **1972**, *72*, 705−718.
(61) Sheather, S. J.; Jones, M. C. A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *J. R. Statist. Soc. B* **1991**, *53*, 683−690.
(62) Scott, D. W. *Multivariate Density Estimation and Visualization*; Humboldt-Universität Berlin, Center for Applied Statistics and Economics (CASE), 2004; pp 1−23.
(63) Burton, G. W.; Sims, L. B.; Wilson, J. C.; Fry, A. Calculation of Carbon-14, Chlorine-37, and Deuterium Kinetic Isotope Effects in the Solvolysis of tert-Butyl Chloride. *J. Am. Chem. Soc.* **1977**, *99*, 3371−3379.
(64) Pauling, L. Atomic Radii and Interatomic Distances in Metals. *J. Am. Chem. Soc.* **1947**, *69*, 542−553.
(65) Addadi, L.; Jaffe, E. K.; Knowles, J. R. Secondary Tritium Isotope Effects as Probes of the Enzymic and Nonenzymic Conversion of Chorismate to Prephenate. *Biochemistry* **1983**, *22*, 4494−4501.
(66) Wright, S.; DeClue, M.; Mandal, A.; Lee, L.; Wiest, O.; Cleland, W.; Hilvert, D. Isotope Effects on the Enzymatic and Nonenzymatic Reactions of Chorismate. *J. Am. Chem. Soc.* **2005**, *127*, 12957−12964.
(67) Pawlak, J. L.; Padykula, R. E.; Kronis, J. D.; Aleksejczyk, R. A.; Berchtold, G. A. Structural Requirements for Catalysis by Chorismate Mutase. *J. Am. Chem. Soc.* **1989**, *111*, 3374−3381.
(68) Kawabata, T. Detection of multiscale pockets on protein surfaces using mathematical morphology. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1195−1211.
(69) Řezáčová, P.; Borek, D.; Moy, S. F.; Joachimiak, A.; Otwinowski, Z. Crystal structure and putative function of small Toprim domain-containing protein from *Bacillus stearothermophilus*. *Proteins* **2008**, *70*, 311−319.

DOI: 10.1021/acs.jpcb.8b11059
J. Phys. Chem. B 2019, 123, 448−456

Supporting Information:
Molecular Determinants for the Rate Acceleration in the Claisen Rearrangement Reactions

Sebastian Brickel [1], Markus Meuwly[1,*]

[1]Department of Chemistry, University of Basel, Klingelbergstrasse 80, Basel, Switzerland

(Dated: December 6, 2018)

S1. COMPUTATIONAL DETAILS

The genetic algorithm utilised for the fitting of the GAPOs was adopted from Reference[1] and is based on Differential Evolution (DE)[2]. DE is a powerful optimisation method which explores regions where the objective function has low values. Required inputs are the reference energies and the potential energies of the MS-ARMD surfaces. The maximum number of generations was set 10000, the population size $20 * N_{\text{pars}}$, where N_{pars} is the number of parameters in the GAPO fit (here 15), the crossover ratio is 1, and the differential weight is also 1. The algorithm is terminated when the maximum number of generations is reached .

In order to validate the stability of the enzyme in the MD simulations the calculations in the enzyme were also performed in a smaller water box but on a longer time scale. The cubic box size was reduced to 86 Å^3. The RMSD(t) of the backbone of BsCM for 500 ps of free dynamics shows for the apo-enzyme, as well as the enzyme with AVE as ligand, a value of ≈ 1 Å. For the simulations of the apo-enzyme in the 86 Å^3 water box the free dynamics was extended to 14 ns. A slow, steady increase, which flattens out after 11 ns, staying below 2 Å was observed. The simulation set up for the calculations was therefore assumed to be stable.

The force field parameters for the reactant state are summarized in Tables S3 to S5 and those for the product state are displayed in Tables S6 to S9.

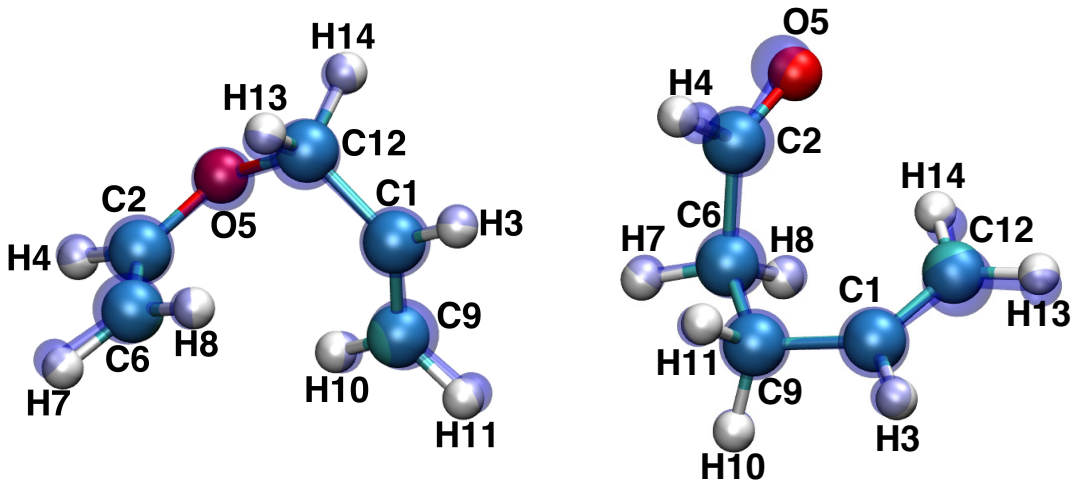


FIG. S1: Alignment between optimised structures using *ab initio* (blue, transparent) and the fitted FF (solid) for the reactant (AVE, left, RMSD = 0.047 Å) and for the product structure (pent-4-enal, right, RMSD = 0.098 Å). Structures include atom numbering and element. Atom color code: C grey, O red, and H white.

MP2 [cm^{-1}]	**FF** [cm^{-1}]	**MP2** [cm^{-1}]	**FF** [cm^{-1}]	**MP2** [cm^{-1}]	**FF** [cm^{-1}]
3311.28	3381.32	1451.64	1554.36	932.54	975.80
3287.80	3377.35	1413.58	1477.85	875.14	897.52
3231.74	3329.31	1357.67	1308.66	806.52	784.38
3207.23	3327.50	1315.86	1293.58	711.34	695.53
3191.61	3289.92	1313.43	1248.92	649.54	677.77
3175.42	3265.55	1235.64	1199.94	567.48	616.17
3131.15	3264.93	1106.63	1086.48	536.37	528.06
3062.57	3215.64	1066.85	1082.06	339.79	377.19
1695.51	1860.05	1019.08	1077.43	288.62	305.92
1669.94	1811.41	1006.36	1025.56	190.49	186.27
1514.93	1654.25	988.75	1020.67	145.48	148.75
1459.27	1578.23	935.18	994.94	102.87	98.71

TABLE S1: Computed vibrational frequencies of the reactant (AVE).

MP2 [cm^{-1}]	**FF** [cm^{-1}]	**MP2** [cm^{-1}]	**FF** [cm^{-1}]	**MP2** [cm^{-1}]	**FF** [cm^{-1}]
3268.58	3550.49	1464.88	1524.05	932.24	827.06
3197.21	3493.24	1432.00	1493.54	915.39	815.89
3167.55	3431.68	1388.40	1358.19	866.35	740.12
3136.69	3083.39	1371.58	1321.21	740.20	657.26
3104.14	3080.39	1316.38	1276.21	680.58	597.36
3082.30	3041.00	1280.85	1197.43	626.19	586.31
3058.84	3016.35	1232.36	1174.97	426.36	500.16
2951.73	3013.79	1159.49	1156.34	325.78	329.83
1745.75	2032.20	1084.16	1130.60	232.42	280.39
1685.95	1729.79	1039.86	1013.14	162.29	149.35
1498.27	1575.95	1037.29	998.08	108.30	132.35
1468.45	1537.41	933.50	855.27	80.92	70.82

TABLE S2: Computed vibrational frequencies of the product (pent-4-enal).

Atom 1 #	Atom 2 #	$k/2$ [kcal/mol/Å^2]	r_e [Å]
1	3	432.89	1.0808
1	9	676.48	1.3344
1	12	315.18	1.4907
2	4	432.89	1.0808
2	5	455.08	1.3538
2	6	676.48	1.3344
6	7	432.89	1.0808
6	8	432.89	1.0808
9	10	432.89	1.0808
9	11	432.89	1.0808
12	13	415.01	1.0912
12	14	415.01	1.0912

Atom 1 #	Atom 2 #	D_e [kcal/mol]	r_e [Å]	β [1/Å]
5	12	96.281	1.4390	1.50740

Atom 1 #	Atom 2 #	Atom 3 #	$k/2$ [kcal/mol/radian^2]	θ_e [degree]
3	1	9	35.7960	124.790
3	1	12	39.9310	120.810
9	1	12	58.0880	123.910
4	2	5	46.7020	112.610
4	2	6	35.7960	124.790
5	2	6	60.9440	124.450
2	5	12	77.9710	110.820
2	6	7	35.7960	124.790
2	6	8	35.7960	124.790
7	6	8	32.6810	124.240
1	9	10	35.7960	124.790
1	9	11	35.7960	124.790
10	9	11	32.6810	124.240
1	12	5	77.0610	114.470
1	12	13	55.9020	115.170
1	12	14	55.9020	115.170
5	12	13	50.0990	110.880
5	12	14	50.0990	110.880
13	12	14	45.9980	113.020

TABLE S3: Bond and angle reactant state parameters.

Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [kcal/mol]	ϕ [degree]
1	12	5	2	3	0.124100	0.00
2	5	12	13	3	0.127380	0.00
2	5	12	14	3	0.127380	0.00
3	1	9	10	2	5.08720	180.00
3	1	9	11	2	5.08720	180.00
3	1	12	5	2	0.525180E-01	180.00
3	1	12	5	3	0.246910	0.00
3	1	12	13	1	-0.325170	0.00
3	1	12	13	2	-0.868130E-01	180.00
3	1	12	13	3	0.129270	0.00
3	1	12	14	1	-0.325170	0.00
3	1	12	14	2	-0.868130E-01	180.00
3	1	12	14	3	0.129270	0.00
4	2	5	12	1	0.800000	0.00
4	2	5	12	2	1.48100	180.00
4	2	5	12	3	0.440400	0.00
4	2	6	7	2	5.08720	180.00
4	2	6	8	2	5.08720	180.00
5	2	6	7	2	5.95060	180.00
5	2	6	8	2	5.95060	180.00
5	12	1	9	1	0.209360	0.00
5	12	1	9	2	0.633890E-01	180.00
5	12	1	9	3	-0.331520	0.00
6	2	5	12	1	-0.803920	0.00
6	2	5	12	2	1.48310	180.00
6	2	5	12	3	-0.399040	0.00
9	1	12	13	1	0.311670	0.00
9	1	12	13	2	-0.155010	180.00
9	1	12	13	3	-0.204300	0.00
9	1	12	14	1	0.311670	0.00
9	1	12	14	2	-0.155010	180.00
9	1	12	14	3	-0.204300	0.00
10	9	1	12	2	5.66970	180.00
11	9	1	12	2	5.66970	180.00
Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [,kcal/mol/radian^2]	ϕ [degree]
1	9	12	3	0	1.15880	0.00000
12	5	1	13	0	0.00000	0.00000
12	5	1	14	0	0.00000	0.00000
2	6	5	4	0	2.29190	0.00000
6	7	2	8	0	0.33063	0.00000
9	10	1	11	0	0.33063	0.00000

TABLE S4: Dihedral and improper reactant state parameters.

Atom #	q_i [e]	$\epsilon_{i,1}$ [kcal/mol]	$R_{min,1}/2$ [Å]	$\epsilon_{i,2}$ [kcal/mol]	$R_{min,2}/2$ [Å]
1	-0.294780	0.513240E-01	1.83800	X	X
2	-0.091549	0.513240E-01	1.83800	X	X
3	0.119000	0.166050E-01	1.43250	X	X
4	0.119000	0.166050E-01	1.43250	X	X
5	-0.289570	0.140120	1.45350	X	X
6	-0.255670	0.513240E-01	1.83800	X	X
7	0.119000	0.166050E-01	1.43250	X	X
8	0.119000	0.166050E-01	1.43250	X	X
9	-0.283650	0.513240E-01	1.83800	X	X
10	0.119000	0.166050E-01	1.43250	X	X
11	0.119000	0.166050E-01	1.43250	X	X
12	0.501210	0.419120E-01	2.20700	0.0124360	1.45780
13	0.000000	0.166050E-01	1.43250	X	X
14	0.000000	0.166050E-01	1.43250	X	X

TABLE S5: Non-bonded reactant state parameters, including special 1-4 CHARMM van der Waals paramters. "X" indicates that this parameter is not needed.

Atom 1 #	Atom 2 #	$k/2$ [kcal/mol/Å^2]	r_e [Å]
1	3	478.07	1.0802
1	9	245.74	1.4868
1	12	545.69	1.3321
2	4	360.60	1.1081
2	5	852.40	1.2184
2	6	285.82	1.5122
6	7	364.54	1.0938
6	8	364.54	1.0938
9	10	364.54	1.0938
9	11	364.54	1.0938
12	13	478.07	1.0802
12	14	478.07	1.0802

Atom 1 #	Atom 2 #	D_e [kcal/mol]	r_e [Å]	β [1/Å]
6	9	53.2620	1.51710	1.58570

Atom 1 #	Atom 2 #	Atom 3 #	$k/2$ [kcal/mol/radian^2]	θ_e [degree]
3	1	9	41.0460	116.760
3	1	12	48.9390	120.150
9	1	12	53.6620	124.010
4	2	5	88.3700	127.950
4	2	6	63.0710	126.590
5	2	6	91.5160	131.710
2	6	7	32.4980	108.360
2	6	8	32.4980	108.360
2	6	9	50.1500	112.460
7	6	8	40.8990	106.940
7	6	9	44.8760	111.100
8	6	9	44.8760	111.100
1	9	6	57.8290	111.740
1	9	10	36.2200	110.380
1	9	11	36.2200	110.380
6	9	10	44.8760	111.100
6	9	11	44.8760	111.100
10	9	11	40.8990	106.940
1	12	13	48.9390	120.150
1	12	14	48.9390	120.150
13	12	14	27.6110	115.790

TABLE S6: Bond and angle product state parameters.

Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [kcal/mol]	ϕ [degree]
1	9	6	2	3	0.166260	0.00
1	9	6	7	1	0.183010	0.00
1	9	6	7	2	-0.391550	180.00
1	9	6	7	3	0.824110E-01	0.00
1	9	6	8	1	0.183010	0.00
1	9	6	8	2	-0.391550	180.00
1	9	6	8	3	0.824110E-01	0.00
2	6	9	10	1	-0.200010	0.00
2	6	9	10	2	0.258020E-01	180.00
2	6	9	11	1	-0.200010	0.00
2	6	9	11	2	0.258020E-01	180.00
3	1	9	6	1	0.283310E-01	0.00
3	1	9	6	3	0.115020	0.00
3	1	9	10	1	-0.483200	0.00
3	1	9	10	2	-0.130700	180.00
3	1	9	10	3	0.416390E-01	0.00
3	1	9	11	1	-0.483200	0.00
3	1	9	11	2	-0.130700	180.00
3	1	9	11	3	0.416390E-01	0.00
3	1	12	13	2	7.17700	180.00
3	1	12	14	2	7.17700	180.00
4	2	6	7	1	-0.379530	0.00
4	2	6	7	2	0.252370	180.00
4	2	6	7	3	0.417030	0.00
4	2	6	8	1	-0.379530	0.00
4	2	6	8	2	0.252370	180.00
4	2	6	8	3	0.417030	0.00
4	2	6	9	1	-0.691640E-01	0.00
4	2	6	9	2	0.326580	180.00
4	2	6	9	3	0.198250	0.00
5	2	6	7	1	0.140450	0.00
5	2	6	7	2	-0.484600	180.00
5	2	6	7	3	0.584940E-01	0.00
5	2	6	8	1	0.140450	0.00
5	2	6	8	2	-0.484600	180.00
5	2	6	8	3	0.584940E-01	0.00
5	2	6	9	1	0.377480	0.00
5	2	6	9	2	0.496460E-01	180.00
5	2	6	9	3	0.689800E-01	0.00
6	9	1	12	1	-0.201670	0.00
6	9	1	12	2	0.890890E-01	180.00
6	9	1	12	3	-0.259520	0.00

TABLE S7: Dihedral product state parameters (part I).

Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [kcal/mol]	ϕ [degree]
7	6	9	10	1	0.973920E-01	0.00
7	6	9	10	2	-1.27110	180.00
7	6	9	10	3	0.140120	0.00
7	6	9	11	1	0.973920E-01	0.00
7	6	9	11	2	-1.27110	180.00
7	6	9	11	3	0.140120	0.00
8	6	9	10	1	0.973920E-01	0.00
8	6	9	10	2	-1.27110	180.00
8	6	9	10	3	0.140120	0.00
8	6	9	11	1	0.973920E-01	0.00
8	6	9	11	2	-1.27110	180.00
8	6	9	11	3	0.140120	0.00
9	1	12	13	2	6.11170	180.00
9	1	12	14	2	6.11170	180.00
10	9	1	12	1	0.108170	0.00
10	9	1	12	2	-0.185050	180.00
10	9	1	12	3	-0.184290	0.00
11	9	1	12	1	0.108170	0.00
11	9	1	12	2	-0.185050	180.00
11	9	1	12	3	-0.184290	0.00
Atom 1 #	Atom 2 #	Atom 3 #	Atom 4 #	N	k [kcal/mol/radian^2]	ϕ [degree]
1	12	9	3	0	1.60150	0.00
9	6	1	10	0	0.0000	0.00
9	6	1	11	0	0.0000	0.00
6	2	9	7	0	0.0000	0.00
6	2	9	8	0	0.0000	0.00
2	5	6	4	0	5.53920	0.00
12	13	1	14	0	0.354580	0.00

TABLE S8: Dihedral (part II) and improper product state parameters.

Atom #	q_i [e]	$\epsilon_{i,1}$ [kcal/mol]	$R_{min,1}/2$ [Å]	$\epsilon_{i,2}$ [kcal/mol]	$R_{min,2}/2$ [Å]
1	-0.162180	0.121160	1.95720	X	X
2	0.195600	0.531170E-01	1.70270	X	X
3	0.114670	0.262580E-01	1.33790	X	X
4	0.052634	0.262580E-01	1.33790	X	X
5	-0.388760	0.181880	1.11690	0.20246	1.24060
6	0.029608	0.280630E-01	1.62660	0.01795	1.16760
7	0.000000	0.262580E-01	1.33790	X	X
8	0.000000	0.262580E-01	1.33790	X	X
9	0.124170	0.280630E-01	1.62660	0.01795	1.16760
10	0.000000	0.262580E-01	1.33790	X	X
11	0.000000	0.262580E-01	1.33790	X	X
12	-0.198330	0.121160	1.95720	X	X
13	0.114670	0.262580E-01	1.33790	X	X
14	0.114670	0.262580E-01	1.33790	X	X

TABLE S9: Non-bonded product state parameters.

The barrier region connecting the reactant and product state is described by three GAPOs

$$\Delta V^{ij}_{\mathrm{GAPO,k}}(x) = \exp\Big(-\frac{(\Delta V_{ij}(x) - V^0_{ij,k})^2}{2\sigma^2_{ij,k}}\Big) \times \sum_{l=0}^{m_{ij,k}} a_{ij,kl}(\Delta V_{ij}(x) - V^0_{ij,k})^l \qquad \text{(S1)}$$

with the parameters summarized in Table S10.

k	$V^0_{ij,k}$	$\sigma_{ij,k}$	$a_{ij,k0}$	$a_{ij,k1}$	$a_{ij,k2}$	$a_{ij,k3}$
1	-79.412387703	126.51811456	-17.145048202	-0.10147938960	-2.9295083192E-04	-6.6958795252E-07
2	161.37689877	259.12300169	378.19011783	-0.37733302902		
3	107.35317360	251.81649573	-426.34677642	0.11825095138		

TABLE S10: GAPO parameters: i labels the reactant, j labels the product, $\mathrm{V}^0_{ij,k}$ is the center of the Gaussian function (in kcal/mol), and $\sigma_{ij,k}$ the width of the Gaussian (in kcal/mol). a_{ij} is the polynomial coefficient in kcal/mol$^{(1-j)}$, $j = 0, 1, 2, 3$

S2. ADDITIONAL ANALYSIS OF THE INDIVIDUAL SYSTEMS

AVE For comparison, mixed semiempirical (DFTB3[3])/molecular mechanics simulations were also performed with the same simulation conditions as for the MS-ARMD simulations. AVE was the QM part and the environment (solvent and enzyme plus solvent, respectively) was treated with molecular mechanics. Semiempirical DFTB3 was chosen since it offers a good balance between cost and efficiency. The activation enthalpy in the gas phase is 19.7 kcal/mol which underestimates the experimentally determined $\Delta H^{\ddagger}$ (30.6 kcal/mol). On the other hand, the gas phase free energy barrier height from US is $\Delta G^{\ddagger} = 31.8$ kcal/mol. These results are in accord with previous simulations which find $\Delta H^{\ddagger} = 19.5$ kcal/mol and $\Delta G^{\ddagger} = 24.2$ kcal/mol using SPC/E water and the selective integrate-over temperature (SITS) method for sampling.[4] With DFTB in aqueous solution the TS is again stabilized ($\Delta\Delta G^{\ddagger}_{\text{(gas-solv)}} = -2.2$ kcal/mol) (but less so) and no stabilization in the enzyme BsCM was observed, see thinner solid lines in top panel of Figure 3.

AVE-CO2: Calculations in the gas phase were also performed with DFTB3 (see middle panel in Figure 5) which yields a barrier height of 36.3 kcal/mol. This confirms the trend of the MS-ARMD simulations ($\Delta G^{\ddagger}_{\text{MS-ARMD}} = 32.2$ kcal/mol) that gas phase barriers for AVE-$(CO_2)_2$ are higher than those for AVE. This is consistent with previous electronic structure calculations at the RHF/6-31G* level of theory.[5]. Experimental $\Delta H^{\ddagger}$ or $\Delta G^{\ddagger}$ values for AVE-$(CO_2)_2$ are not available. In aqueous solution calculated with MS-ARMD the $\Delta\Delta G^{\ddagger}_{\text{gas-solv}}$ is 17 kcal/mol, more than twice the value for AVE.

Comparing the results for AVE in BsCM with those obtained from gas phase simulations both methods, MS-ARMD and DFTB3, yield a stabilisation by 1.6 kcal/mol. In comparison to the results obtained in aqueous solution no further stabilisation was found in BsCM ($\Delta\Delta G^{\ddagger}_{\text{solv-enz}} = +4.6$ kcal/mol). The barrier height from MS-ARMD for the AVE to pent-4-enal reaction in BsCM in a 86 Å^3 water box yields a barrier height of 24.9 kcal/mol compared with 23.9 kcal/mol for simulations in the larger (116 Å^3) box. A further difference between the two water box sizes around BsCM calculated with MS-ARMD can be seen in the $r_c^{\ddagger}$ position. In the smaller water box the transition state is shifted towards the product side (-0.48 Å compared with -0.66 Å). Notably, the $r_c^{\ddagger}$ position in the 86 Å^3 water

box is even further shifted to larger positive values of r_c than for the simulations in gas phase. This is contrary to the effect displayed in solution and enzyme, where the position of the TS moves further to the reactant side compared to gas phase. For both box sizes the maximum of the PMF is located at a higher value than for the results obtained in solution.

Chorismate: The gas-phase simulations of chorismate, (green line in the bottom panel of Figure 3) gives a $\Delta G^{\ddagger}$ of 20.5 kcal/mol at $r_c^{\ddagger} = -0.99$ Å. This barrier height is between the results obtained from AVE and AVE-$(CO_2)_2$. Experimental $\Delta G^{\ddagger}$ values for the Claisen rearrangement of Chorismate in gas phase are not available.

Computationally, comparing the gas phase result to aqueous solution a stabilisation of 3.7 kcal/mol is found. The experimental $\Delta G^{\ddagger}$ is 24.5 kcal/mol[6], compared with 16.8 from simulations (see bottom panel in Figure 3). The $r_c^{\ddagger}$ from MS-ARMD calculations in solution moves closer to the reactant side, as for AVE and AVE-$(CO_2)_2$. The free energy of the product was found to be considerably lower compared to expectation, see blue and green solid lines. Therefore the simulation time of the dynamics was increased to 1 ns for each window. The $\Delta G^{\ddagger}$ value changes only slightly with the increased simulation time (16.0 kcal/mol compared to 16.8 kcal/mol for 50 ps of simulations per window) compared to the experimental $\Delta G^{\ddagger} = 24.5$ kcal/mol[6]. However the product is destabilized by 6 kcal/mol. This suggests that in the current US simulations the product is not equilibrated and longer simulation times should lead to further increase of the energy of the product.

S3. MORE O'FERRALL-JENCKS ANALYSIS

In order to clarify the mechanistic details involved in the rearrangement "More O'Ferrall-Jencks" (MOFJ) diagrams[7,8] were generated for all US simulations (50 ps per window). Such diagrams involve two progression coordinates and are useful to illustrate their coupling during the course of a reaction. The corners of the diagram correspond to the reactant, product and the two non-concerted extremes respectively[9], see left panel in Figure S3.

To test the influence of the value of the constant C in the Pauling relation, MOFJ diagrams

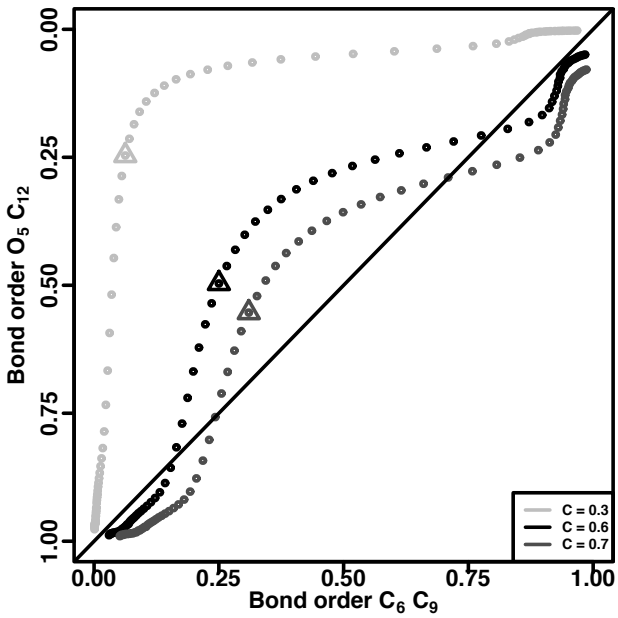


FIG. S2: More O'Ferrall-Jencks diagrams for MP2 IRC of the conversion of AVE to pent-4-enal. Points show the structures from the IRC, triangles mark the TS structure. Bond orders are defined using the interatomic distance for the bonds from the respective *ab initio* optimised structure in the Pauling relation (equation 2).

have been generated for a range of values for the MP2 IRC, see Figure S2. A value of 0.71 Å was proposed by Pauling[10] for describing the bond order between C-C single, double, and triple bonds. $C = 0.6$ Å has been suggested for long bonds[11] and $C = 0.3$ Å was defined by revision of the Pauling relation[12].

Snapshots from simulations of a downhill simulations for chorismate in gas phase from the TS towards either product or reactant were also analysed such that they can be represented as a More O'Ferrall-Jencks diagram. This was done in order to verify that the bias in the US simulations does not artificially alter the sampling and that the More O'Ferrall-Jencks diagrams from US are meaningful. As Figure 5 shows, the data from the unbiased downhill simulations (black dots) show different path than the contours from the US simulations (red). Especially around the TS the distribution of the C6-C9 bond order is shifted. This is an entropic effect, since the reaction coordinate chosen for the US does not specify the C-C and C-O distance specifically.

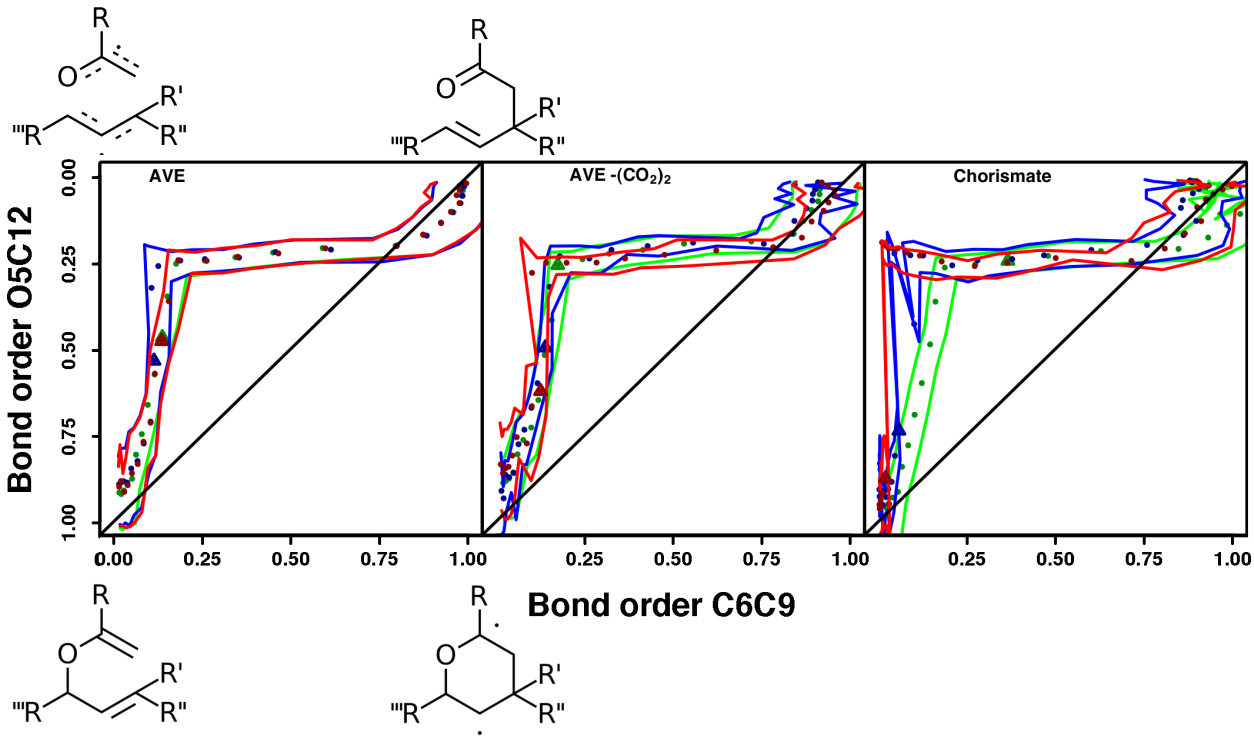


FIG. S3: More O'Ferrall-Jencks diagrams for Claisen rearrangement of AVE (left), AVE-2,6-dicarboxylate (middle) and of chorismate (right) for simulations in gas phase (green), solvent (blue), and in the enzyme BsCM (red). Points represent the arithmetic average of the distribution from the 50 ps US trajectories, lines the maximum and minimum from KDE. Bond orders are defined using the interatomic distance for the bonds from the respective *ab initio* optimised structure in the Pauling relations (equation 2).

A. Center of mass distributions

The normalised probability distribution of $x-$, $y-$ and $z-$coordinates of the CoM of the heavy atom framework of the three different molecules (equivalent to heavy atoms in AVE, see atom numbering in Figure 1A) within the active site was analysed for the reactant (Figure S4) and product (Figure S5) conformations. This provides insight into spatial effects on the reaction.

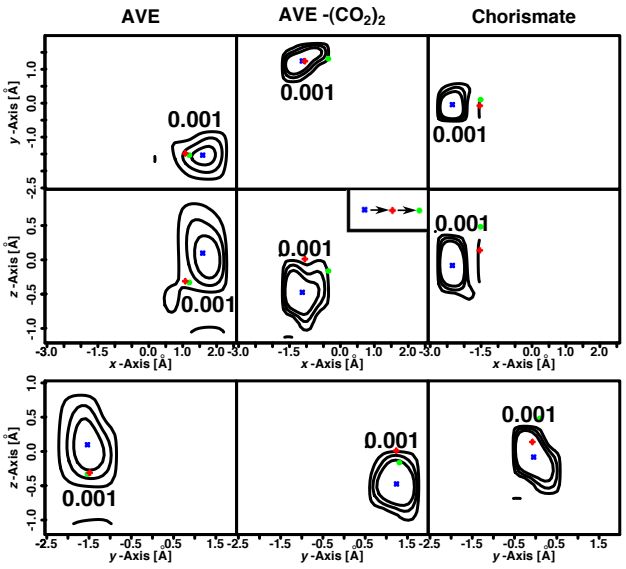


FIG. S4: Normalised probability distribution of the $x-$, $y-$, and $z-$coordinate of the center of mass of the heavy atom framework of AVE (left), AVE-$(CO_2)_2$ (middle), and chorismate (right) from the first umbrella, i.e. the reactant structure. The center of gravity of the reactant (blue x), TS (red plus), and product (green dot) are displayed in each panel.

The positioning of AVE and its product is displaced by 2 to 3 Å along the $y-$ and $x-$axes, respectively, compared to AVE-$(CO_2)_2$ and Chorismate and in general all distributions are considerably broadened (see Figures S4, S5 and 6). For AVE-$(CO_2)_2$ in the reactant state the distribution along the $z-$axis is displaced by 0.5 Å compared to the TS (Figures S4 and 6). The center of mass of the distribution for chorismate along the $x-$axis is moved further along the negative $x-$axis which may be caused by steric effects due to the ring compared to AVE-$(CO_2)_2$. To substantiate this the residues in contact with the substrate were analysed. Consistent with the literature[13] it was found that Arg9, Phe57, Cys75, Arg90, and Tyr108 are in direct contact with AVE-$(CO_2)_2$ and Chorismate. Due to its increased size, Chorismate makes additional contacts with Ala59 (through chorismate's -OH group) and Leu345 compared to AVE-$(CO_2)_2$ which contributes to further reduce the barrier.

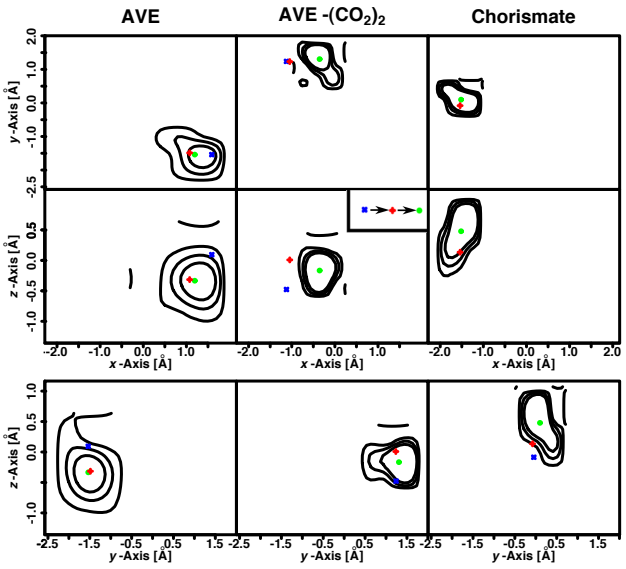


FIG. S5: Normalised probability distribution of the $x-$, $y-$, and $z-$coordinate of the center of mass of the heavy atom framework of AVE (left), AVE-$(CO_2)_2$ (middle), and chorismate (right) from the last umbrella, i.e. the product structure. The center of gravity of the reactant (blue x), TS (red plus), and product (green dot) are displayed in each panel.

[1] Unke, O. T.; Devereux, M.; Meuwly, M. *The Journal of Chemical Physics* **2017**, *147*, 161712.

[2] Li, H.; Nichols, P. G. H.; Han, S.; Foster, K. J.; Sivasithamparam, K.; Barbetti, M. J. *Australas. Plant Pathol.* **2009**, *38*, 284–287.

[3] Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569–585.

[4] Zhang, J.; Yang, Y. I.; Yang, L.; Gao, Y. Q. *J. Phys. Chem. B* **2015**, *119*, 14505–14514.

[5] Wiest, O.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 11628–11639.

[6] Andrews, P. R.; Smith, G. D.; Young, I. G. *Biochemistry* **1973**, *12*, 3492–3498.

[7] O'Ferrall, R. A. M. *J Chem Soc* **1970**, #, 274–277.

[8] Jencks, W. P. *Chem. Rev.* **1972**, *72*, 705–718.

[9] Gajewski, J.; Conrad, N. *J. Am. Chem. Soc.* **1979**, *101*, 6693–6704.

[10] Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ihaca, New York, 1960.

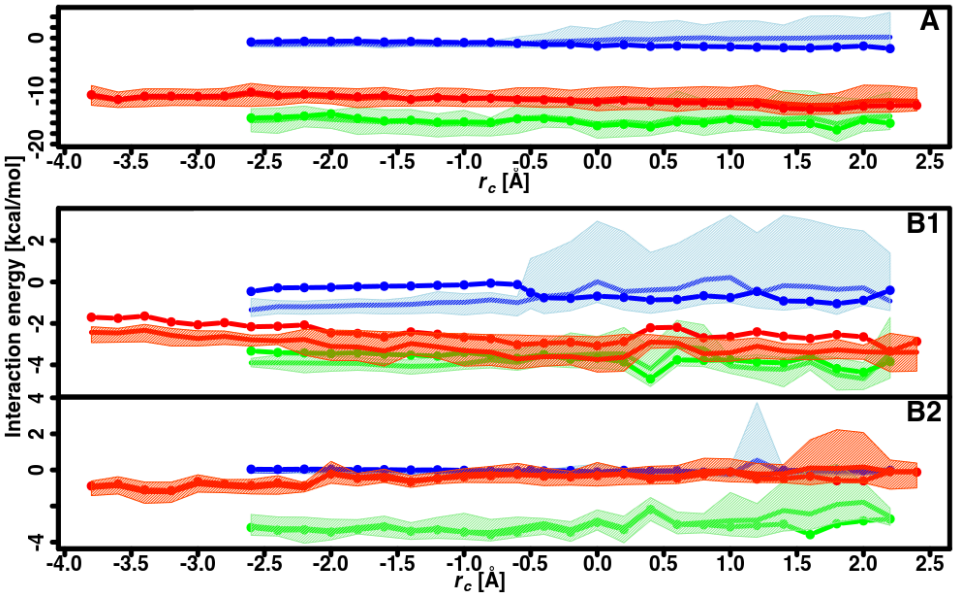
[11] Meyer, M. P.; DelMonte, A. J.; Singleton, D. A. *J. Am. Chem. Soc.* **1999**, *121*, 10865–10874.

[12] Burton, G. W.; Sims, L. B.; Wilson, J. C.; Fry, A. *J. Am. Chem. Soc.* **1977**, *99*, 3371–3379.

[13] Burschowsky, D.; van Eerde, A.; Ökvist, M.; Kienhöfer, A.; Kast, P.; Hilvert, D.; Krengel, U. *Proc. Natl. Acad. Sci.* **2014**, *111*, 17516–17521.

Part IV.

Conclusion



No great thing is created suddenly.

- *Epictetus* -

6. General Conclusion and Outlook

Reactive molecular dynamics simulations allow for a refined picture of chemical reaction mechanism on an atomistic level, both in gas and condensed phase. The reactive MD method MS-ARMD was shown to be a versatile tool for a broad range of chemical reactions. Its basis, the combination of force fields with a priority on the energetically lowest one and switching between the surfaces by an energy dependent switching function, enables it to be as capable as QM/MM simulations. Its capability of calculating statistically significant number of trajectories for sufficiently long time (up to converged reaction rates) allows for a more in depth investigation of mechanistic details, from small molecules in gas phase to enzymatic catalysis.

Vibrationally induced OH-stretch dynamics on thousands of trajectories was found to lead to elimination for HSO_3Cl when 3 to 6 quanta are excited. The process occurs on the sub-ns to ns time scale and it was found that fast dissociating trajectories follow in general the direct dissociation path (without prior H-transfer). For longer time after OH-stretch excitation, indirect dissociation becomes the dominant process. The two different pathways for the fragmentation reaction yield a characteristic energy distribution for HCl that can be experimentally validated. Furthermore, a normal mode decomposition scheme in combination with the development and application of the concept of a MDP shows the importance of the "umbrella" mode (SO_3 bend) for the dissociation to occur. In contrast to that, HSO_3F does not show significant fragmentation, even after excitation with more than 6 quanta of OH-stretch excitation energy. H-transfer occurs starting with $\nu_{OH} > 4$. The initial H-transfer is limited by the energy flow into the necessary reaction coordinate and the expectation time shows a broad distribution, highlighting the dependence on the initial conditions. Subsequent H-transfer residence times decay exponentially. Extensive sampling of the dissociation reaction reveals a dynamic bottleneck in comparison to HSO_3Cl due to the coupling between angular and radial degrees of freedom, that hinders the dissociation in HSO_3F. Reaction rates for the unimolecular fragmentation reaction of the

two studied systems, as well as for HSO_3Br show a clear trend connected to the commonly known trends in the periodic table, regarding size and electronegativity.

MS-ARMD parametrisations were found to be rather general, e.g. they can easily be transferred from gas to aqueous solution. For the S_N2 reaction of $[Br\text{-}CH_3 + Cl]^-$ the gas phase simulations reproduced the experimental values within chemical accuracy and the PMF displayed a similar behaviour to a neural network fitted to thousands of reference structures. The same parametrisation in aqueous solvent overestimated the experimentally determined barrier height. This was traced back to the utilised charge model. Nevertheless the PMF followed the expected behaviour, i.e. the disappearance of energy increase for large separation of the ion from the molecules as displayed in the gas phase. This separates MS-ARMD from the other reactive FF method tested on this system (MS-VALBOND) where this gas phase feature of the PMF is still visible in aqueous solution. Compared to other reactive molecular dynamics methods the effort required for the parametrisation of MS-ARMD (e.g. number of required reference structures) was found to be moderate (in between MS-VALBOND and the neural network).

The fitting of the reactive FF can still be considered a bottleneck in all reactive force field methods as it is labour intensive. However, depending on the goal in mind, it is possible to transfer a given parametrisation to other molecules which undergo the same reaction, given that the substituents (with respect to the fitted molecule) do not actively participate in the reaction. This was shown for the Claisen rearrangement reaction, where the goal was the systematic assessment of the reaction with three different molecules (AVE, AVE-$(CO_2)_2$, and chorismate) in three different environments (gas phase, aqueous solution, and enzyme CM). Even though specific parametrisation for each molecule in the particular environment could improve the quantitative agreement with experiment, it was found that, for the Claisen rearrangement, utilizing the reactive parametrisation of AVE for the biological relevant molecule chorismate does lead to the expected outcome of a continuous speed up when going from gas phase over aqueous solution to the enzyme. This is in line with transition state theory and further underpins the substrate specificity of enzymes, as the other two molecules do not display a rate acceleration when going from aqueous solution to the enzyme. The reaction mechanism was found to be step wise with the bond breaking preceding the bond formation. This effect becomes more pronounced when going from gas to condensed phase, especially in the presence of the enzyme. The extensive sampling of MS-ARMD further allowed to analyse the positioning of the molecules structures within

the enzymes active site. Overall the catalytic power of the enzyme could be traced back, thanks to direct comparison of the three molecules, to entropic contribution, as well as electrostatic contributions from the charged carboxylate groups. This can help to design new inhibitors for this enzyme.

Further work related to enzymatic catalysis will include methylation reaction with S-Adenosyl methionine (SAM) and phosphorylation reaction. Furthermore the Claisen rearrangement can also be studied in different solvents, such as methanol and benzene. Studying further methyl transfer reaction, e.g. the Menshutkin reaction as well as the solvent effects on this reaction.

Futur work utilising MS-ARMD will include studies of different methyl transfer reactions, such as the Menshutkin reaction, where the reactant state is neutral, in contrast to the S_N2 where it is ionic. Solvent effects on the rate coefficients of these Menshutkin reactions, as well as on the Claisen rearrangement can provide further insight in the reaction mechanism. Moreover such studies can aid in explaining the different catalytic activities of different solvents, exeeding the simple picture of polar vs. apolar solvents. Studies of different enzymatic reactions, like on methylation and phosphorilation, can be used to increase the understanding of these essential biochemical reactions. On the technical side futur improvements for MS-ARMD might inlcude paralellization of the code to achieve a higher performence, especially in condensed phase. Furthermore the inclusion of different exponential decay paramters for the individual reactions will increase the flexibility of MS-ARMD when describing reaction networks.

In summary, the utilization of reactive MD simulations contributes significantly to a broad range of chemical investigations, by providing molecular level insight into reaction mechanisms on significant time scales. The incorporation of MS-ARMD within the CHARMM software package allows for a widespread use.

Bibliography

[1] P. R. Andrews, G. D. Smith and I. G. Young, *Biochemistry*, 1973, **12**, 3492–3498.

[2] X. Ji, X. Song, J. Li, Y. Bai, W. Yang and X. Peng, *J. Am. Chem. Soc.*, 2007, **129**, 13939–48.

[3] P. Kast, M. Asif-Ullah and D. Hilvert, *Tetrahedron Lett.*, 1996, **37**, 2691–2694.

[4] K. El Hage, S. Brickel, S. Hermelin, G. Gaulier, L. Bonacina, S. C. V. Keulen, S. Bhattacharyya, M. Chergui, P. Hamm, U. Rothlisberger, J.-P. Wolf and M. Meuwly, *Struct. Dyn.*, 2017, **4**, 061507.

[5] M. Meuwly, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2018, e1386.

[6] A. K. Das and M. Meuwly, *Angew. Chemie Int. Ed.*, 2018, **57**, 3509–3513.

[7] P. R. Brooks, *Chem. Rev.*, 1988, **88**, 407–428.

[8] T. Nagy and M. Meuwly, in *Theory and Applications of the Empirical Valence Bond Approach*, ed. F. Duarte and L. Kamerlin, John Wiley & Sons Ltd., 2017, ch. 1, pp. 1–25.

[9] B. R. Brooks, C. L. Brooks III, A. D. Mackerell Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545–1614.

[10] H. M. Senn and W. Thiel, *Angew. Chem. Int. Ed. Engl.*, 2009, **48**, 1198–229.

[11] A. R. Lacey, *J. Chem. Educ.*, 1987, **64**, 756–761.

[12] M. J. Ramos and P. A. Fernandes, *Acc. Chem. Res.*, 2008, **41**, 689–698.

[13] A. Warshel and R. M. Weiss, *J. Am. Chem. Soc.*, 1980, **102**, 6218–6226.

[14] M. H. Schmid, A. K. Das, C. R. Landis and M. Meuwly, *J Chem Theory Comput.*, 2018, **15**, 3565–3578.

[15] T. Nagy, J. Yosa Reyes and M. Meuwly, *J. Chem. Theory Comput.*, 2014, **10**, 1366–1375.

[16] D. R. Glowacki, A. J. Orr-Ewing and J. N. Harvey, *J. Chem. Phys.*, 2015, **143**, 044120.

[17] J. Florián, M. F. Goodman and A. Warshel, *J. Am. Chem. Soc.*, 2003, **125**, 8163–8177.

[18] M. Roca, A. Vardi-Kilshtain and A. Warshel, *Biochemistry*, 2009, **48**, 3046–3056.

[19] M. Štrajbl, A. Shurki, M. Kato and A. Warshel, *J. Am. Chem. Soc.*, 2003, **125**, 10228–10237.

[20] A. Yadav, R. M. Jackson, J. J. Holbrook and A. Warshel, *J. Am. Chem. Soc.*, 1991, **113**, 4800–4805.

[21] M. Fuxreiter and A. Warshel, *J. Am. Chem. Soc.*, 1998, **120**, 183–194.

[22] B. Hartke and S. Grimme, *Phys. Chem. Chem. Phys.*, 2015, **17**, 16715–16718.

[23] J. K. Hwang and A. Warshel, *J. Am. Chem. Soc.*, 1996, **118**, 11745–11751.

[24] S. Brickel, A. K. Das, O. T. Unke, H. T. Turan and M. Meuwly, *Electronic Structure*, 2019, **1**, 024002.

[25] J. Yosa Reyes, T. Nagy and M. Meuwly, *Phys. Chem. Chem. Phys.*, 2014, **16**, 18533–18544.

[26] S. Brickel and M. Meuwly, *J. Phys. Chem. A*, 2017, **121**, 5079–5087.

[27] J. Yosa Reyes, S. Brickel, O. T. Unke, T. Nagy and M. Meuwly, *Phys. Chem. Chem. Phys.*, 2016, **18**, 6780–6788.

[28] W. Heitler and F. London, *Z. Phys.*, 1927, **44**, 455–472.

[29] R. N. Porter and M. Karplus, *J. Chem. Phys.*, 1964, **40**, 1105–1115.

[30] R. M. Van Doren, *J. Chem. Educ.*, 2012, **89**, 584–584.

[31] F. Duarte, A. Pabis and S. C. L. Kamerlin, in *Introduction to the Empirical Valence Bond Approach*, John Wiley & Sons, Ltd, 2017, ch. 2, pp. 27–61.

[32] S. Brickel and M. Meuwly, *J. Phys. Chem. B*, 2019, **123**, 448–456.

[33] M. Soloviov, A. K. Das and M. Meuwly, *Angew. Chem. Int. Ed.*, 2016, **55**, 10126–10130.

[34] L. Pauling, *J. Am. Chem. Soc.*, 1947, **69**, 542–553.

[35] A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard III, *J. Phys. Chem. A*, 2001, **41**, 9396–9409.

[36] P. O. Hubin, D. Jacquemin, L. Leherte and D. P. Vercauteren, *J. Comput. Chem.*, 2016, **37**, 2564–2572.

[37] M. Raju, P. Ganesh, P. R. C. Kent and A. C. van Duin, *J. Chem. Theory Comput.*, 2015, **11**, 2156–2166.

[38] B. J. Finlayson-Pitts, *Anal. Chem.*, 2010, **82**, 770–776.

[39] V. Vaida, K. J. Feierabend, N. Rontu and K. Takahashi, *Int. J. Photoenergy*, 2008, **2008**, 1–13.

[40] V. Vaida and D. J. Donaldson, *Phys. Chem. Chem. Phys.*, 2014, **16**, 827–36.

[41] D. J. Donaldson, A. F. Tuck and V. Vaida, *Chem. Rev.*, 2003, **103**, 4717–4729.

[42] D. J. Hofmann and J. M. Rosen, *Nature*, 1982, **297**, 120–124.

[43] P. Rasch, S. Tilmes, R. Turco, A. Robock, L. Oman, C.-C. Chen, G. Stenchikov and R. Garcia, *Philos. Trans. R. Soc. Lond., A*, 2008, **306**, 4007–4037.

[44] F. F. Crim, *Annu. Rev. Phys. Chem.*, 1984, **35**, 657–691.

[45] F. F. Crim, *J. Phys. Chem.*, 1996, **100**, 12725–12734.

[46] F. F. Crim, *Proc. Nat. Acad. Sci. USA*, 2008, **105**, 12654–12661.

[47] P. M. Hundt, B. Jiang, M. E. van Reijzen, H. Guo and R. D. Beck, *Science*, 2014, **344**, 504–507.

[48] J. Wei, B. Karpichev and H. Reisler, *J. Chem. Phys.*, 2006, **125**, 034303.

[49] E. Kamarchik, C. Rodrigo, J. M. Bowman, H. Reisler and A. I. Krylov, *J. Chem. Phys.*, 2012, **136**, 084305.

[50] L. Fono, D. Donaldson, R. J. Proos and B. R. Henry, *Chem. Phys. Lett.*, 1999, **311**, 131–138.

[51] R. B. Gerber and J. Sebek, *Int. Rev. Phys. Chem.*, 2009, **28**, 207–222.

[52] C. M. Roehl, S. A. Nizkorodov, H. Zhang, G. A. Blake and P. O. Wennberg, *J. Phys. Chem. A*, 2002, **106**, 3766–3772.

[53] Y. Miller, G. M. Chaban and R. B. Gerber, *Chem. Phys.*, 2005, **313**, 213–224.

[54] M. Staikova, D. J. Donaldson and J. S. Francisco, *J. Phys. Chem. A*, 2002, **106**, 3023–3028.

[55] Y. Miller, G. M. Chaban and R. B. Gerber, *J. Phys. Chem. A*, 2006, 5342–5354.

[56] H. Zhang, C. M. Roehl, S. P. Sander and P. O. Wennberg, *J. Geophys. Res.*, 2000, **105**, 14593–14598.

[57] J. Matthews, R. Sharma and a. Sinha, *J. Phys. Chem. A*, 2004, **108**, 8134–8139.

[58] K. Takahashi, K. L. Plath, R. T. Skodje and V. Vaida, *J. Phys. Chem. A*, 2008, **112**, 7321–7331.

[59] K. L. Plath, K. Takahashi, R. T. Skodje and V. Vaida, *J. Phys. Chem. A*, 2009, **113**, 7294–7303.

[60] G. da Silva, *J. Phys. Chem. A*, 2016, **120**, 276–283.

[61] L. M. Reynard and D. J. Donaldson, *J. Phys. Chem. A*, 2002, **106**, 8651–8657.

[62] V. Vaida, H. G. Kjaergaard, P. E. Hintze and D. J. Donaldson, *Science*, 2003, **299**, 1566–1568.

[63] F. Raes, R. V. Dingenen, E. Vignati, J. Wilson, J.-P. Pataud, J. H. Seinfeld and P. Adams, *Atmos Environ A-Gen*, 2000, **34**, 4215–4240.

[64] U. Pöschl, *Angew. Chem. Int. Ed. Engl.*, 2005, **44**, 7520–7540.

[65] D. J. Hofmann and J. M. Rosen, *Geophys. Res. Lett.*, 1985, **12**, 13–16.

[66] J. Zhao, O. B. Toon and R. P. Turco, *J. Geophys. Res., D*, 1995, **100**, 5215–5227.

[67] C. P. Rinsland, M. R. Gunson, M. K. W. Ko, D. W. Weisenstein, R. Zander, M. C. Abrams, A. Goldman, N. D. Sze and G. K. Yue, *Geophys. Res. Lett.*, 1995, **22**, 1109–1112.

[68] M. J. Mills, O. B. Toon and S. Solomon, *Geophys. Res. Lett.*, 1999, **26**, 1133–1136.

[69] J. B. Burkholder, M. Mills and S. McKeen, *Geophys. Res. Lett.*, 2000, **27**, 2493–2496.

[70] P. E. Hintze, H. G. Kjaergaard, V. Vaida and J. B. Burkholder, *J. Chem. Phys.*, 2003, **107**, 1112–1118.

[71] S. J. Wrenn, L. J. Butler, G. A. Rowland, C. J. Knox and L. F. Phillips, *J. Chem. Phys.*, 1999, **129**, 1112–1118.

[72] H. G. Kjaergaard, J. R. Lane, A. L. Garden, D. P. Schofield, T. W. Robinson and M. Mills, *Adv. Quantum Chem.*, 2008, **55**, 1112–1118.

[73] T. W. Robinson, D. P. Schofield and H. G. Kjaergaard, *J. Chem. Phys.*, 2003, **118**, 7226–7232.

[74] K. J. Feierabend, D. K. Havey, S. S. Brown and V. Vaida, *Chem. Phys. Lett.*, 2006, **420**, 438–442.

[75] J. Yosa Reyes and M. Meuwly, *J. Phys. Chem. A*, 2011, **115**, 14350–14360.

[76] C. A. Rivera, N. Winter, R. V. Harper, I. Benjamin and S. E. Bradforth, *Phys. Chem. Chem. Phys.*, 2011, **13**, 8269–8283.

[77] A. J. Orr-Ewing, *J. Chem. Phys.*, 2014, **140**, 090901.

[78] S. Shaik, *Prog. Phys. Org. Chem.*, 1985, **15**, 197–337.

[79] J. Chandrasekhar, S. F. Smith and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1985, **107**, 154–163.

[80] A. Merkel, R. Zahradník and Z. Havlas, *J. Am. Chem. Soc.*, 1988, **110**, 8355–8359.

[81] J. Gao and X. Xia, *J. Am. Chem. Soc.*, 1993, **115**, 9667–9675.

[82] X. Fradera, L. Amat, M. Torrent, J. Mestres, P. Constans, E. Besald, J. Marti, S. Simon, M. Lobato, J. M. Oliva, J. M. Luis, M. Sol, R. Carbb and M. Duran, *J. Mol. Struct.*, 1996, **371**, 171–183.

[83] I. Adamovic and M. S. Gordon, *J. Phys. Chem. A*, 2005, **109**, 1629–1636.

[84] S. Shaik, A. Ioffe, A. C. Reddy and A. Pross, *J. Am. Chem. Soc.*, 2006, **116**, 262–273.

[85] D. L. Severance and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1992, **114**, 10966–10968.

[86] J. M. Guest, J. S. Craw, M. A. Vincent and I. H. Hillier, *J. Chem. Soc. Perkin Trans. 2*, 1997, 71–74.

[87] C. J. Cramer and D. G. Truhlar, *J Am Chem Soc*, 1992, **114**, 8794–8799.

[88] A. C. Moskun and S. E. Bradforth, *J. Chem. Phys.*, 2003, **119**, 4500–4515.

[89] S. Martí, M. Roca, J. Andrés, V. Moliner and E. Silla, *Chem. Soc. Rev.*, 2004, **33**, 98–107.

[90] A. Lodola and A. J. Mulholland, *Methods Mol. Biol.*, 2013, **924**, 67–89.

[91] J.-M. Choi, S.-S. Han and H.-S. Kim, *Biotechnol. Adv.*, 2015, **33**, 1443 – 1454.

[92] G. L. Butterfoss and B. Kuhlman, *Annu. Rev. Biophys.*, 2006, **35**, 49–65.

[93] R. J. Pantazes, M. J. Grisewood and C. D. Maranas, *Curr. Opin. Struct. Biol.*, 2011, **21**, 467 – 472.

[94] M. W. Van der Kamp and A. J. Mulholland, *Biochemistry*, 2013, **52**, 2708–2728.

[95] S. J. Benkovic and S. Hammes-Schiffer, *Science*, 2003, **301**, 1196–1202.

[96] A. K. Das and M. Meuwly, *Empirical Force Fields for Mechanistic Studies of Chemical Reactions in Proteins*, Elsevier Inc., 1st edn., 2016, vol. 577, pp. 31–55.

[97] S. Mishra and M. Meuwly, *J. Am. Chem. Soc.*, 2010, **132**, 2968–2982.

[98] A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12562–12566.

[99] J. Kirkwood, *J. Chem. Phys.*, 1935, 300–313.

[100] M. Meli and G. Colombo, *Int. J. Mol. Sci.*, 2013, **14**, 12157–12169.

[101] J. Kottalam and D. A. Case, *J. Am. Chem. Soc.*, 1988, **110**, 7690–7697.

[102] D. A. McQuarrie and J. D. Simon, *Physical Chemistry : a Molecular Approach*, University Science Books, Sausalito, Calif., 1997.

[103] M. Born and R. Oppenheimer, *Annalen der Physik*, 1927, **389**, 457–484.

[104] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, Inc., Mineola, 1st edn., 1996.

[105] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Inc., USA, 2006.

[106] O. Christiansen, H. Koch and P. Jørgensen, *Chem. Phys. Lett.*, 1995, **243**, 409–418.

[107] C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618–622.

[108] D. G. Tomlinson, A. Asadchev and M. S. Gordon, *J. Comput. Chem.*, 2016, **37**, 1274–1282.

[109] R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724–728.

[110] T. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.

[111] Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim and M. Karplus, *J. Phys. Chem. B*, 2001, **105**, 569–585.

[112] J. Pu, J. Gao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 5454–5463.

[113] W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, 1133–1138.

[114] W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657–1666.

[115] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.

[116] R. Kosloff, *J. Phys. Chem.*, 1988, **92**, 2087–2100.

[117] W. Yu, X. He, K. Vanommeslaeghe and A. D. Mackerell Jr., *J Comput Chem.*, 2012, **33**, 2451–2468.

[118] P.-A. Cazade, H. Tran, T. Bereau, A. K. Das, F. Kläsi, P. Hamm and M. Meuwly, *J. Chem. Phys.*, 2015, **142**, 212415.

[119] S. Jo, T. Kim, V. G. Iyer and W. Im, *J. Comput. Chem.*, 2008, **29**, 1859–1865.

[120] S. Kim, J. Lee, S. Jo, C. L. Brooks, H. S. Lee and W. Im, *J. Comput. Chem.*, 2017, **38**, 1879–1886.

[121] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant,

S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision D.01*, Gaussian Inc. Wallingford CT 2009.

[122] J. Nelder and R. Mead, *Chem. Phys.*, 1965, **7**, 308–313.

[123] J. Lee, H. A. Scheraga and S. Rackovsky, *J. Comput. Chem.*, 1997, **18**, 1222–1232.

[124] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J. Comput. Chem.*, 1998, **19**, 1639–1662.

[125] H. Li, P. G. H. Nichols, S. Han, K. J. Foster, K. Sivasithamparam and M. J. Barbetti, *Australas. Plant Pathol.*, 2009, **38**, 284–287.

[126] O. T. Unke, M. Devereux and M. Meuwly, *J. Chem. Phys.*, 2017, **147**, 161712.

[127] V. Zoete, M. A. Cuendet, A. Grosdidier and O. Michielin, *J. Chem. Comp.*, 2011, **32**, 2959–2368.

[128] L. Verlet, *Phys. Rev.*, 1967, **159**, 98–103.

[129] J. R. Lane, H. G. Kjaergaard, K. L. Plath and V. Vaida, *J. Phys. Chem. A*, 2007, **111**, 5434–5440.

[130] P. Gupta, J. R. Lane and H. G. Kjaergaard, *Phys. Chem. Chem. Phys.*, 2010, **12**, 8277–8284.

[131] J. R. Lane and H. G. Kjaergaard, *J. Phys. Chem. A*, 2007, **111**, 9707–9713.

[132] O. T. Unke, S. Brickel and M. Meuwly, *J. Chem. Phys.*, 2019, **150**, 074107.

[133] J. R. Barker, T. L. Nguyen, J. F. Stanton, C. Aieta, M. Ceotto, F. Gabas, T. J. D. Kumar, C. G. L. Li, L. L. Lohr, A. Maranzana, N. F. Ortiz, J. M. Preses, J. M. Simmie, J. A. Sonk and P. J. Stimac, *MultiWell-2017 Software Suite*, MultiWell-2017 Software Suite, 1991.

[134] S. Gronert, C. H. DePuy and V. M. Bierbaum, *J. Am. Chem. Soc.*, 1991, **113**, 4009–4010.

[135] W. L. Hase, *Science*, 1994, **266**, 998–1002.

[136] C. A. Lieder and J. I. Brauman, *Int. J. Mass Spectrom. Ion Phys.*, 1975, **16**, 307–319.

[137] W. N. Olmstead and J. I. Brauman, *J. Am. Chem. Soc.*, 1977, **99**, 4219–4228.

[138] P. M. Hierl, A. F. Ahrens, M. Henchman, A. A. Viggiano and J. F. Paulson, *Int. J. Mass Spectrom. Ion Process.*, 1987, **81**, 101–122.

[139] S. E. Barlow, J. M. Van Doren and V. M. Bierbaum, *J. Am. Chem. Soc.*, 1988, **110**, 7240–7242.

[140] S. R. Vande Linde and W. L. Hase, *J. Am. Chem. Soc.*, 1989, **111**, 2349–2351.

[141] C. H. DePuy, S. Gronert, A. Mullin and V. M. Bierbaum, *J. Am. Chem. Soc.*, 1990, **112**, 8650–8655.

[142] S. Kato, G. E. Davico, H. S. Lee, C. H. DePuy and V. M. Bierbaum, *Int. J. Mass Spectrom.*, 2001, **210-211**, 223–229.

[143] M. Stei, E. Carrascosa, M. A. Kainz, A. H. Kelkar, J. Meyer, I. Szabó, G. Czakó and R. Wester, *Nat. Chem.*, 2016, **8**, 151–156.

[144] E. Carrascosa, J. Meyer, J. Zhang, M. Stei, T. Michaelsen, W. L. Hase, L. Yang and R. Wester, *Nat. Commun.*, 2017, **8**, 1–7.

[145] O. T. Unke and M. Meuwly, *J. Chem. Theory. Comput.*, 2019, **15**, 3678–3693.

[146] S. Schmatz, P. Botschwina and H. Stoll, *Int. J. Mass Spectrom.*, 2000, **201**, 277–282.

[147] G. Purvis and R. Bartlett, *J. Chem. Phys.*, 1982, **76**, 1910–1918.

[148] A. D. MacKerell Jr., D. Bashford, M. Bellott, R. L. Dunbrack Jr., J. D. Evenseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.

[149] D. Bates and D. Watts, *Nonlinear Regression Analysis and its Applications*, Wiley, New York, 1988.

[150] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013.

[151] W. L. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey and M. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.

[152] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Compt. Chem.*, 1992, **13**, 1011–1021.

[153] B. Bogdanov and T. McMahon, *Int. J. Mass Spectrom.*, 2005, **241**, 205–223.

[154] R. H. Bathgate and E. A. Moelwyn-Hughes, *J. Chem. Soc.*, 1959, 2642–2648.

[155] S. W. Sigsworth and A. W. Castleman, *J. Am. Chem. Soc.*, 1992, **114**, 10471–10477.

[156] J. Peng, N. Puskas, P. B. Corkum, D. M. Rayner and A. V. Loboda, *Anal. Chem.*, 2012, **84**, 5633–5640.

[157] L. Pedraza-González, J. F. Galindo, R. González and A. Reyes, *J. Phys. Chem. A*, 2016, **120**, 8360–8368.

[158] R. M. Coates, B. D. Rogers, S. J. Hobbs, D. P. Curran and D. R. Peck, *J. Am. Chem. Soc.*, 1987, **35**, 2601–2605.

[159] L. Claisen, *Chem. Ber.*, 1912, **45**, 3157–3166.

[160] W. N. White and E. F. Wolfarth, *J. Org. Chem.*, 1970, **35**, 2196–2199.

[161] O. Acevedo and K. Armacost, *J. Am. Chem. Soc.*, 2010, **132**, 1966–1975.

[162] B. Ganem, *Tetrahedron*, 1978, **34**, 3353–3383.

[163] W. P. Jencks, *Chem. Rev.*, 1972, **72**, 705–718.

Part V.

Appendix

Abbreviation and Units

aug-cc-pVXZ	augmented cc-pVXZ
cc-pVXZ	correlation-consistent polarised Valence X-tuble Zeta
CCSD	Coupled-Cluster with Single and Double
CCSD(T)	Coupled-Cluster with Single, Double and perturbative Triple excitations
CGenFF	CHARMM Generalized Force Field
CHARMM	Chemistry at HARvard Molecular Mechanics
CI	Configuration Interaction
CM	Chorismate Mutase
DFTB	Density-Functional Tight-Binding
FES	Free Energy Surface
FF	Force Field
GA	Genetic Algorithm
GLJ	Generalized Lennard Jones
GTO	Gaussian Type Orbital
HF	Hartree-Fock
IRC	Intrinsic Reaction Coordinate
IVR	Intramolecular Vibrational Relaxation
LJ	Lennard-Jones
MD	Molecular Dynamics
MDP	Minimum Dynamic Path
MEP	Minimum Energy Path
MM	Molecular Mechanics
MP2	second order Møller-Plesset perturbation theory

MS-ARMD	Multi-Surface Adiabatic Reactive Molecular Dynamics
MS-VALBOND	Multi-State VALence BOND
NPT	isothermal-isobaric ensemble
NVE	microcanonical ensemble
NVT	canonical ensemble
occ.	occupied orbitals
PES	Potential Energy Surface
PMF	Potential of Mean Force
PT	Perturbation Theory
QM	Quantum Mechanics
RRKM	Rice-Ramsperger-Kassel-Markus
SAM	S-Adenosyl methionine
SCF	Self-Consistent Field
STO	Slater Type Orbital
US	Umbrella Sampling
UV	Ultraviolet
vdW	van der Waals
virt.	virtual orbitals
VXZ	Valence X-tuble Zeta

Å	Ångström
cm^{-1}	Wavenumber
K	Kelvin
kcal	Kilo-calorie
mol	Mole
fs	Femtosecond
ps	Picosecond
ns	Nanosecond